Docket #PF-0126-1 DIV2

"Express Mail" mailing label number EM 062 546 944 US I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Assistant Commissioner for Patents, Box PATENT APPLICATION, Washington, D.C. 20231 on 21-Sept.

By: _____ Printed: John J. Cherry

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
REQUEST FOR FILING A PATENT APPLICATION UNDER 37 CFR 1.53(b)

Assistant Commissioner for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

Dear Sir:

This is a request for filing a **DIVISIONAL** application under 37 CFR 1.53(b) of pending prior application Serial No. **08 /715,204** , filed on **September 18, 1996**, entitled **NOVEL TUMOR PROTEINS**.

1. X Enclosed is a copy of the prior application, U. S. Application Serial No. 08/715,204 filed September 17, 1996 , including the oath or declaration as originally filed.

2. X With regard to the requirement of 37 CFR 1.821(d) which requires that a copy of the Sequence Listing in computer readable form (CRF) be submitted, Applicants submit that the computer readable form of the "Sequence Listing" in the instant divisional/continuation application, is identical with that filed for Serial No. 08/715,204 , filed September 17, 1996 , to which priority is claimed. In accordance with 37 C.F.R. §1.821(e), please use the computer readable form filed with U.S. Application Serial No. 08/715,204 as the computer readable form for the instant divisional application. It is understood that the Patent and Trademark Office will make the necessary change in application number and filing date for the computer readable form that will be used for the instant divisional application.

3. ___ A verified statement to establish small entity status under 37 CFR 1.9 and 1.27 was filed in prior application number 08/_____ and such status is still proper and desired (37 CFR 1.28(a)).

4. X Cancel in this application original claims 1-12 and 14-19 of the prior application before calculating the filing fee. (At least one original independent claim must be retained for filing purposes.)

5. X The inventor(s) of the invention being claimed in this application is (are): Olga Bandman, Janice Au-Young, Surya K. Goli, Jennifer L. Hillman.

6. X In accordance with 37 CFR 1.63(d) a copy of the originally signed declaration showing applicants' signature(s) as filed on January 6, 1997 is enclosed.

7. _X_ Amend the specification by inserting before the first line the sentence: "This application is a divisional application of U.S. application serial number 08/715,204 , filed September 17, 1996."

8. _X_ The filing fee is calculated below:

Claims	Number Filed	Minus	Number Extra	Other Than Small Entity Rate	Fee	Basic Fee $790.00
Total Claims	6	-20	0	x $22	0	$ 0
Indep. Claims	2	-3	0	x $82	0	$0
__ Multiple Dependent Claim(s), if any				+ $270		$ 0

TOTAL FILING FEE $ 790.00

9. __ An extension of time in the above-named prior application has been requested and the fees therefor have been authorized in said application.

10. _X_ Please charge Deposit Amount No. **09-0108** in the amount of $ 790.00 .

The Commissioner is hereby authorized to charge any additional fees which may be required or credit any overpayment to Account No. **09-0108**.

Duplicate copies of these papers are enclosed.

11. _X_ New formal drawings are enclosed.

12. __ Priority of foreign application number _____ , filed on _____ in _____ is claimed under 35 U.S.C. 119.

13. _X_ The prior application is assigned of record to Incyte Pharmaceuticals, Inc.

14. _X_ A preliminary amendment is enclosed.

15. _X_ Also enclosed: Information Disclosure Statement
List of References Cited by Applicants
Copy of an Associate Power of Attorney filed May 15, 1998

16. _X_ The power of attorney of the prior application is to:

LUCY J. BILLINGS **Registration No. 36,749**
MICHAEL C. CERRONE **Registration No. 39,132**

a. ____ The power of attorney appears in the original papers in the prior application.

b. _X_ A copy of the associate power in the prior application is enclosed.

c. _X_ An executed Associate Power of Attorney is enclosed.

c. _X_ Address all future correspondence to: (May only be completed by applicant, or attorney or agent of record.)

Legal Department
INCYTE PHARMACEUTICALS, INC.
PATENT DEPARTMENT
3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555, Fax: (650) 845-4166

Date: 29 September 98 By: Leanne C. Price

Leanne C. Price
Reg. No. 42,090

___ Inventor(s)
___ Assignee of complete interest
___ Attorney or agent of record
_X_ Filed under 37 CFR 1.34(a)
Registration number if acting under 37 CFR 1.34(a) 42,090.

"Express Mail" mailing label number EM 062 546 944 US I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Assistant Commissioner for Patents, Box PATENT APPLICATION, Washington, D.C. 20231 on 29 Sept 98

By: _____

Printed: John J. Cherry

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Bandman et al.

Title:	A NOVEL TUMOR PROTEIN		
Serial No.:	To Be Assigned	Filing Date:	Herewith
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

Applicants request reconsideration of the above-referenced patent application in view of the following amendments and remarks.

IN THE SPECIFICATION

Please amend the specification as follows:

At page 1, line 5, replace "prevention" with --prevention,--;

At page 2, line 21, replace "as shown in" with --shown in the--;

At page 2, line 22, replace "sequence" with --sequences--;

At page 2, line 40, replace "Figures 1A and 1B" with --Figures 1A, 1B, and 1C--;

At page 3, line 1, replace "MacDNAsis" with --MACDNASIS--;

At page 3, line 3, replace "Figures 2A and 2B" with --Figures 2A, 2B, and 2C--;

At page 3, line 6, replace "LIFESEQ™" with --the LIFESEQ®--;

At page 3, line 8, replace "Figure 4 shows" with --Figure 4A and 4B show--;

At page 3, line 12, replace "DNAStar" with --DNASTAR--;

At page 3, line 14, replace "NO:3)," with --NO:3)--;

At page 3, line 16, replace "NO:1; the" with --NO:1. The--;

At page 3, line 17, replace "negative Y axis," with --Y axis reflects--;

At page 3, line 17, replace "Figs. 7-10" with --Figures 7, 8, 9, and 10--;

At page 3, line 35, replace "GCG Fragment Assembly System" with --GCG fragment assembly system--;

At page 4, line 4, replace "inserted" with --inserted,--;

At page 4, line 16, replace "DNAStar" with --DNASTAR--;

At page 5, line 19, replace "Figs. 3 and 4" with --Figures 3, 4A, and 4B--;

At page 5, line 20, replace "Fig." with --Figure--;

At page 6, line 17, replace "Figures 1A, 1B, 2A, and 2B" with --Figures 1A, 1B, 1C, 2A, 2B, and 2C--;

At page 7, line 16, replace "Figures 1A, 1B, 2A, and 2B" with --Figures 1A, 1B, 1C, 2A, 2B, and 2C--;

At page 8, line 6, delete "Sequenase®" and insert --SEQUENASE--;

At page 8, line 10, replace "Amplification System" with --amplification system--;

At page 8, line 11, replace "Micro Lab" with --MICROLAB--;

At page 8, line 12, replace "Thermal Cycler" with --thermal cycler--;

At page 9, line 5, replace "PromoterFinder™ Clontech (Palo Alto CA)" with --PROMOTERFINDER (Clontech, Palo Alto, CA)--

At page 9, line 20, delete "devise" and insert --device--;

At page 9, line 21, delete "Genotyper" and insert --GENOTYPER--;

At page 9, line 22, delete "Sequence Navigator™" and insert --SEQUENCE NAVIGATOR--;

At page 11, line 30, delete "Bluescript®" and insert --BLUESCRIPT--;

At page 12, line 5, delete "Bluescript®" and insert --BLUESCRIPT phagemid--;

At page 17, line 12, replace "Figs. 3 and 4" with --Figures 3, 4A, and 4B--;

At page 29, line 3, replace "ratio, LD50/ED50" with --ratio ED50/LD50--;

At page 30, line 4, replace "Uni-ZAP™" with --UNIZAP--;

At page 30, line 11, replace "UniZAP" with --UNIZAP--;

At page 30, line 13, replace "pBluescript®" with --pBLUESCRIPT--;

At page 30, line 17, replace "E. Coli host strain XL1-Blue®" with --XL1-BLUE E.coli host strain--;

At page 30, lines 23 through 25, delete "(lot #0024A; Mayo Clinic, Rochester MN) by retropubic prostatectomy. The pathology report indicated Mayo grade 3 of 4 adenocarcinoma (Gleason grade 3+3)";

At page 30, line 31, replace "Homogenizer Polytron-PT 3000" with --Polytron-PT 3000 homogenizer--;

At page 30, line 39, replace "Qiagen Oligotex" with --OLIGOTEX--;

At page 31, line 7, replace "LambdaZap®" with --LAMBDAZAP--;

At page 31, line 8, replace "pBluescript™" with --pBLUESCRIPT--;

At page 31, line 9, replace "XL1-BlueMRF™" with --XL1-BLUEMRF--;

At page 31, line 26, replace "pBluescript®" with --pBLUESCRIPT--;

At page 31, line 31, replace "QIAwell-8 Plasmid Purification" with --QIAWELL-8 plasmid purification--;

At page 31, line 32, replace "System™ from Qiagen®, QIAwell PLUS™, or QIAwell ULTRA™ DNA Purification" with --system or the QIAWELL PLUS or QIAWELL ULTRA DNA purification--;

At page 31, line 33, replace "Systems" with --systems--;

At page 31, line 38, replace "Micro Lab" with --MICROLAB--;

At page 31, line 39, replace "Thermal Cyclers" with --thermal cyclers--;

At page 31, line 40, replace "Sequencing Systems" with --sequencing systems--;

At page 32, line 3, replace "INHERIT™" with --INHERIT--;

At page 32, line 4, replace "Sequence Analysis System" with --sequence analysis system--;

At page 32, line 15, replace "INHERIT- 670 Sequence Analysis System" with --INHERIT-670 sequence analysis system--;

At page 33, line 8, replace "LIFESEQ™ database (Incyte, Palo Alto CA)" with --LIFESEQ® database (Incyte Pharmacueticals, Inc., Palo Alto, CA)--;

At page 33, line 38, delete "U.S. Patent Application 08/487,112, filed June 7, 1995, specifically incorporated by reference";

At page 33, line 40, replace "OLIGO® 4.06 Primer" with --OLIGO 4.06 primer--;

At page 34, line 1, replace "Analysis Software" with --analysis software--;

At page 34, line 14, replace "Thermal Cycler" with --thermal cycler--;

At page 34, line 35, delete "QIAQuick™" and insert --QIAQUICK--;

At page 35, line 43, replace "AR™" with --AR--;

At page 36, line 1, replace "blots" with --blots, or the blots are exposed--;

At page 36, line 2, delete "for several hours";

At page 36, line 11, delete "Figs. 1A, 1B, 2A, and 2B" and insert --Figures 1A, 1B, 1C, 2A, 2B, and 2C--;

At page 36, line 13, delete "Figures 1A, 1B, 2A, and 2B," and insert --Figures 1A, 1B, 1C, 2A, 2B, and 2C--;

At page 36, line 20, delete "Figures 1A, 1B, 2A, and 2B" and insert --Figures 1A, 1B, 1C, 2A, 2B, and 2C--;

At page 37, line 17, replace "DNAStar" with --DNASTAR--;

At page 37, line 21, replace "Figs." with --Figures--;

At page 37, line 23, replace "Peptide Synthesizer Model 431 A" with --431A peptide synthesizer--; and

At page 37, line 36, replace "Sepharose" with --SEPHAROSE--.

IN THE CLAIMS

Please cancel claims 1-12 and 14-19 without prejudice.

REMARKS

Justification for the amendments is as follows. The amendments to the specification were made to correct mere typographical or grammatical errors, to reflect the renumbering of the figures in the preparation of formal drawings, or to delete portions of the sepcification. In the claims, claims 1-12 and 14-19 were canceled, and claims 13 and 20-24 are thus pending. No new matter is added by any of these amendments.

If a telephone conference would expedite the prosecution of the present application, or if the Examiner has any questions, the Examiner is invited to call the undersigned (650) 845-4672.

Respectfully submitted,

INCYTE PHARMACEUTICALS, INC.

Date: 29 September 98

Leanne C. Price
Reg. No. 42,090

3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 845-4672
Fax: (650) 845-4166

PF 0126US

NOVEL TUMOR PROTEINS

FIELD OF THE INVENTION

The present invention relates to nucleic acid and amino acid sequences of novel human tumor proteins and to the use of these sequences in the diagnosis, study, prevention and treatment of disease.

BACKGROUND OF THE INVENTION

In an effort to identify genes involved in the early stages of tumor progression, Bryne JA et al (1995, Cancer Res 55: 2896-2903) screened human breast and basal cell carcinomas for genes commonly overexpressed in tumor cells relative to non-tumor tissue. One novel sequence, D52, was differentially expressed in carcinoma cells and showed little homology to other genes. Chen SL et al (1996, Oncogene 12: 741-751) independently cloned D52 based on its increased expression in lung tumor derived cell lines relative to cell lines derived from normal tissues. Recently, Byrne et al (1996, Genomics 35: 523-532) described a human homolog of D52, termed D53, which is often coexpressed with D52 and may form hetero- or homo-dimers. Both D52 and D53 contain PEST domains, regions that are rich in amino acid residues proline (P), glutamate (E), serine (S), and threonine (T) (Rechsteiner M (1990) Semin Cell Biol 1: 433-440). In human D52, 18 of 37 amino terminal residues are PEST domain residues (Byrne et al, supra). PEST domains are associated with rapidly degraded enzymes, transcriptional factors, and components of receptor signaling pathways (Loetscher P et al (1991) J Biol Chem 266: 11213-11220). Caenorhabditis elegans open reading frame (ORF) F13E6.1 has homology to D52 (Wilson R et al (1994) Nature 368: 32-38).

Tumor Proteins and Disease

Cancer remains a major public health concern, and current preventative measures and treatments do not match the needs of most patients. For example, among women in the United States, as many as one in eight will contract breast cancer in their lifetime (Helzlsouer KJ (1994) Curr Opin Oncol 6: 541-548). Furthermore, the incidence of breast cancer is rising by about 1% a year (Harris JR et al (1992) N Engl J Med 327: 319-328). Among men over 50 years of age, the lifetime risk of prostate cancer is 9.5% and of death from prostate cancer is 2.9% (McLellan DL et al (1995) Can Med Assoc J 153: 895-900).

Genes may be differentially expressed in tumor cells relative to non-tumor cells. For example, elevated expression levels of 12-lipooxygenase

correlate with advanced stage and poor differentiation of human prostate cancer (Gao X et al (1995) Urology 46: 227-237). Additionally, the high incidence of HER2 gene overexpression in breast tumors suggests that perturbations in HER2 are among the earliest and most common genetic lesions in human breast cancer (Liu E et al (1992) Oncogene 7: 1027-1032). This correlation has led to the development of potential HER2 specific therapeutics (Kern JA et al (1993) Am J Respir Cell Mol Biol).

The discovery of additional tumor associated genes may provide agents which are more efficacious in cancer diagnosis and treatment than HER2 and 12-lipooxygenase. Novel tumor associated genes may be specific to a different spectrum of tumor types than known genes. A new tumor protein would satisfy a significant need in the art by providing new agents for the diagnosis, prevention, and treatment of cancer.

SUMMARY

The present invention discloses two novel human tumor proteins (hereinafter referred to individually as TUPROA and TUPROB, and collectively as TUPRO), characterized as having homology to human D52 (GI 790225). and C. elegans ORF ZK418.5 (GI 470373), respectively. Accordingly, the invention features substantially purified tumor proteins, as shown in amino acid sequence of SEQ ID NO:1 and SEQ ID NO:3, and having characteristics of tumor proteins.

One aspect of the invention features isolated and substantially purified polynucleotides which encode TUPRO. In a particular aspect, the polynucleotides are the nucleotide sequences of SEQ ID NO:2 and SEQ ID NO:4. In addition, the invention features polynucleotide sequences that hybridize under stringent conditions to SEQ ID NO:2 or SEQ ID NO:4.

The invention further relates to nucleic acid sequences encoding TUPRO, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof, methods for producing TUPRO or fragments thereof, and use of the sequences in expression vectors and host cells comprising polynucleotides which encode TUPRO. The present invention also relates to antibodies which bind specifically to TUPRO and pharmaceutical compositions comprising substantially purified TUPRO or fragments thereof, or antagonists of TUPRO.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A and 1B show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of the novel tumor protein, TUPROA. The

alignment was produced using MacDNAsis software (Hitachi Software Engineering Co Ltd, San Bruno CA).

Figures 2A and 2B show the amino acid sequence (SEQ ID NO:3) and nucleic acid sequence (SEQ ID NO:4) of the novel tumor protein, TUPROB.

Figure 3 shows the northern analysis for SEQ ID NO:2. The northern analysis was produced electronically using LIFESEQ™ database (Incyte Pharmaceuticals, Palo Alto CA).

Figure 4 shows the northern analysis for SEQ ID NO:4.

Figure 5 shows the amino acid sequence alignments among TUPROA (SEQ ID NO:1), human D52 (GI 790225; SEQ ID NO:5), and C. elegans F13E6.1 (GI 1072344; SEQ ID NO:6). The alignment was produced using the multisequence alignment program of DNAStar software (DNAStar Inc, Madison WI).

Figure 6 shows the amino acid sequence alignments between TUPROB (SEQ ID NO:3), and C. elegans ORF ZK418.5 (GI 470373; SEQ ID NO:6).

Figure 7 shows the hydrophobicity plot (generated using MacDNAsis software) for TUPROA, SEQ ID NO:1; the X axis reflects amino acid position, and the negative Y axis, hydrophobicity (Figs. 7-10).

Figure 8 shows the hydrophobicity plot for human D52, SEQ ID NO:5.

Figure 9 shows the hydrophobicity plot for TUPROB, SEQ ID NO:3.

Figure 10 shows the hydrophobicity plot for C. elegans ORF ZK418.5, SEQ ID NO:7.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, amino acid sequence as used herein refers to oligopeptide or protein sequence.

"Consensus" as used herein may refer to a nucleic acid sequence 1) which has been resequenced to resolve uncalled bases, 2) which has been extended using XL-PCR (Perkin Elmer) in the 5' or the 3' direction and resequenced, 3) which has been assembled from the overlapping sequences of more than one Incyte clone GCG Fragment Assembly System, (GCG, Madison WI), or 4) which has been both extended and assembled.

"Peptide nucleic acid" as used herein refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary (template) strand of nucleic acid (Nielsen PE et al (1993)

Anticancer Drug Des 8:53-63).

As used herein, TUPRO refers to the amino acid sequences of substantially purified TUPRO obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic or recombinant.

A "variant" of TUPRO is defined as an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, eg, replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, eg, replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNAStar software.

A "deletion" is defined as a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

An "insertion" or "addition" is that change in an amino acid or nucleotide sequence which has resulted in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring TUPRO.

A "substitution" results from the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

The term "biologically active" refers to a TUPRO having structural, regulatory or biochemical functions of a naturally occurring TUPRO. Likewise, "immunologically active" defines the capability of the natural, recombinant or synthetic TUPRO, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "derivative" as used herein refers to the chemical modification of a nucleic acid encoding TUPRO or the encoded TUPRO. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of natural TUPRO.

As used herein, the term "substantially purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they

are naturally associated.

"Stringency" typically occurs in a range from about Tm-5°C (5°C below the Tm of the probe)to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, a stringency hybridization can be used to identify or detect identical polynucleotide sequences or to identify or detect similar or related polynucleotide sequences.

The term "hybridization" as used herein shall include "any process by which a strand of nucleic acid joins with a complementary strand through base pairing" (Coombs J (1994) Dictionary of Biotechnology, Stockton Press, New York NY). Amplification as carried out in the polymerase chain reaction technologies is described in Dieffenbach CW and GS Dveksler (1995, PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview NY).

Preferred Embodiments

The present invention relates to novel human tumor proteins and to the use of the nucleic acid and amino acid sequences in the study, diagnosis, prevention and treatment of disease. The sequences encoding TUPRO were found in cDNA libraries from a variety of tissues including several types of tumors (Figs. 3 and 4). TUPROB expression is strongly associated with cDNA libraries derived from tumor tissue (Fig. 4).

The present invention also encompasses TUPRO variants. A preferred TUPRO variant is one having at least 80% amino acid sequence similarity to the TUPRO amino acid sequence (SEQ ID NO:1 or SEQ ID NO:3), a more preferred TUPRO variant is one having at least 90% amino acid sequence similarity to SEQ ID NO:1 or SEQ ID NO:3, and a most preferred TUPRO variant is one having at least 95% amino acid sequence similarity to SEQ ID NO:1 or SEQ ID NO:3.

Nucleic acids encoding the human tumor protein TUPROA of the present invention were first identified in cDNA, Incyte Clone 146723 from a cDNA library made from peripheral blood mononuclear cells, TLYMNOR01, through a computer-generated search for amino acid sequence alignments. The following Incyte clones (and cDNA libraries from which they were derived) were extended and assembled to create the consensus sequence (SEQ ID NO:2): 146723 (TLYMNOR01); 763607 (BRAITUT02); and 601581 (BRSTNOT02). TUPROA, SEQ ID NO:1, is encoded by the nucleic acid sequence of SEQ ID NO:2.

TUPROB was first identified in cDNA, Incyte clone 717832 from a cDNA library made from prostate tumor tissue, PROSTUT01. The following Incyte clones (and cDNA libraries from which they were derived) were extended and assembled to create the consensus sequence (SEQ ID NO:4): 717832 (PROSTUT01); 274790 (PANCDIT03); 628576 (KIDNOT05); 890214 (STOMTUT01); 985743 (LVENNOT03); 1321834 (BLADNOT04); 1398242 (BRAITUT08); and 1733437 (BRSTTUT08). TUPROB, SEQ ID NO:3, is encoded by the nucleic acid sequence

of SEQ ID NO:4.

The present invention is based, in part, on the chemical and structural homology among TUPROA, human D52 (GI 790225; Byrne et al, supra), and C. elegans F13E6.1 (GI 1072344; Wilson et al, supra; Fig. 5). The present invention is also based, in part, on the chemical and structural homology between TUPROB and C. elegans ORF ZK418.5 (GI 470373; Wilson et al, supra) Fig. 6). The novel TUPROA is 204 amino acids long and shares 52% identity with human D52. The novel TUPROB is 245 amino acids long and shares 40% identity with C. elegans ORF ZK418.5. As illustrated by Figures 7 and 8, TUPROA and human D52 have similar hydrophobicity plots suggesting similar structure. TUPROB and C. elegans ORF ZK418.5 have similar hydrophobicity plots suggesting membrane localization (Figs. 9 and 10). TUPROA and TUPROB each have one potential N-glycosylation site.

The TUPRO Coding Sequences

The nucleic acid and deduced amino acid sequences of TUPROA and TUPROB are shown in Figures 1A, 1B, 2A, and 2B. In accordance with the invention, any nucleic acid sequence which encodes the amino acid sequence of TUPRO can be used to generate recombinant molecules which express TUPRO. In a specific embodiment described herein, a nucleotide sequence encoding a portion of TUPROA was first isolated as Incyte Clone 146723 from a peripheral blood mononuclear cell cDNA library (TLYMNOR01). In another specific embodiment described herein, a nucleotide sequence encoding a portion of TUPROB was first isolated as Incyte Clone 717832 from a cDNA library made from prostate tumor tissue, PROSTUT01.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of TUPRO-encoding nucleotide sequences, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene may be produced. The invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring TUPRO, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode TUPRO and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring TUPRO under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TUPRO or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in

accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TUPRO and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

It is now possible to produce a DNA sequence, or portions thereof, encoding a TUPRO and its derivatives entirely by synthetic chemistry, after which the synthetic gene may be inserted into any of the many available DNA vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TUPRO or any portion thereof.

Also included within the scope of the present invention are polynucleotide sequences that are capable of hybridizing to the nucleotide sequence of Figures 1A, 1B, 2A, and 2B under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Berger and Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol 152, Academic Press, San Diego CA) incorporated herein by reference, and may be used at a defined stringency.

Altered nucleic acid sequences encoding TUPRO which may be used in accordance with the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent TUPRO. The protein may also show deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TUPRO. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of TUPRO is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine; glycine, alanine; asparagine, glutamine; serine, threonine phenylalanine, and tyrosine.

Included within the scope of the present invention are alleles of TUPRO. As used herein, an "allele" or "allelic sequence" is an alternative form of TUPRO. Alleles result from a mutation, ie, a change in the nucleic acid sequence, and generally produce altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes which give rise

to alleles are generally ascribed to natural deletions, additions or substitutions of amino acids. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Methods for DNA sequencing are well known in the art and employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland OH)), Taq polymerase (Perkin Elmer, Norwalk CT), thermostable T7 polymerase (Amersham, Chicago IL), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the ABI 377 DNA sequencers (Perkin Elmer).

Extending the Polynucleotide Sequence

The polynucleotide sequence encoding TUPRO may be extended utilizing partial nucleotide sequence and various methods known in the art to detect upstream sequences such as promoters and regulatory elements. Gobinda et al (1993; PCR Methods Applic 2:318-22) disclose "restriction-site" polymerase chain reaction (PCR) as a direct method which uses universal primers to retrieve unknown sequence adjacent to a known locus. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia T et al (1988) Nucleic Acids Res 16:8186). The primers may be designed using OLIGO® 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom M et al (1991) PCR Methods Applic 1:111-19) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before

PCR.

Another method which may be used to retrieve unknown sequences is that of Parker JD et al (1991; Nucleic Acids Res 19:3055-60). Additionally, one can use PCR, nested primers and PromoterFinder libraries to walk in genomic DNA (PromoterFinder™ Clontech (Palo Alto CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes. A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

Capillary electrophoresis may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. Systems for rapid sequencing are available from Perkin Elmer, Beckman Instruments (Fullerton CA), and other companies. Capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled devise camera. Output/light intensity is converted to electrical signal using appropriate software (eg. Genotyper™ and Sequence Navigator™ from Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display is computer controlled. Capillary electrophoresis is particularly suited to the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample. The reproducible sequencing of up to 350 bp of M13 phage DNA in 30 min has been reported (Ruiz-Martinez MC et al (1993) Anal Chem 65:2851-2858).

Expression of the Nucleotide Sequence

In accordance with the present invention, polynucleotide sequences which encode TUPRO, fragments of the polypeptide, fusion proteins or functional equivalents thereof may be used in recombinant DNA molecules that direct the expression of TUPRO in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence, may be used to clone and express TUPRO. As will be understood by those of skill in the art, it may be advantageous to produce TUPRO-encoding nucleotide sequences possessing non-naturally occurring codons. Codons preferred by a particular prokaryotic or eukaryotic host (Murray E et al (1989) Nuc Acids Res 17:477-508) can be selected, for example, to increase the rate of TUPRO

expression or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered in order to alter a TUPRO coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the gene product. For example, mutations may be introduced using techniques which are well known in the art, eg, site-directed mutagenesis to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to produce splice variants, etc.

In another embodiment of the invention, a natural, modified or recombinant polynucleotides encoding TUPRO may be ligated to a heterologous sequence to encode a fusion protein. For example, for screening of peptide libraries for inhibitors of TUPRO activity, it may be useful to encode a chimeric TUPRO protein that is recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between a TUPRO sequence and the heterologous protein sequence, so that the TUPRO may be cleaved and purified away from the heterologous moiety.

In an alternate embodiment of the invention, the coding sequence of TUPRO may be synthesized, whole or in part, using chemical methods well known in the art (see Caruthers MH et al (1980) Nuc Acids Res Symp Ser 215-23, Horn T et al(1980) Nuc Acids Res Symp Ser 225-32, etc). Alternatively, the protein itself could be produced using chemical methods to synthesize a TUPRO amino acid sequence, whole or in part. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge JY et al (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer.

The newly synthesized peptide can be substantially by preparative high performance liquid chromatography (eg, Creighton (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co, New York NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (eg, the Edman degradation procedure; Creighton, supra). Additionally the amino acid sequence of TUPRO, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

Expression Systems

In order to express a biologically active TUPRO, the nucleotide

sequence encoding TUPRO or its functional equivalent, is inserted into an appropriate expression vector, ie, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing a TUPRO coding sequence and appropriate transcriptional or translational controls. These methods include in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination or genetic recombination. Such techniques are described in Sambrook et al (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY and Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY.

A variety of expression vector/host systems may be utilized to contain and express a TUPRO coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (eg, baculovirus); plant cell systems transfected with virus expression vectors (eg, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with bacterial expression vectors (eg, Ti or pBR322 plasmid); or animal cell systems.

The "control elements" or "regulatory sequences" of these systems vary in their strength and specificities and are those nontranslated regions of the vector, enhancers, promoters, and 3' untranslated regions, which interact with host cellular proteins to carry out transcription and translation. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla CA) or pSport1 (Gibco BRL) and ptrp-lac hybrids and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (eg, heat shock, RUBISCO; and storage protein genes) or from plant viruses (eg, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from the mammalian genes or from mammalian viruses are most appropriate. If it is necessary to generate a cell line that contains multiple copies of TUPRO, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for TUPRO. For example, when large

quantities of TUPRO are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the TUPRO coding sequence may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke & Schuster (1989) J Biol Chem 264:5503-5509); and the like. pGEX vectors (Promega, Madison WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems are designed to include heparin, thrombin or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used. For reviews, see Ausubel et al (supra) and Grant et al (1987) Methods in Enzymology 153:516-544.

In cases where plant expression vectors are used, the expression of a sequence encoding TUPRO may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV (Brisson et al (1984) Nature 310:511-514) may be used alone or in combination with the omega leader sequence from TMV (Takamatsu et al (1987) EMBO J 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi et al (1984) EMBO J 3:1671-1680; Broglie et al (1984) Science 224:838-843); or heat shock promoters (Winter J and Sinibaldi RM (1991) Results Probl Cell Differ 17:85-105) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S or Murry LE in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill New York NY, pp 191-196 or Weissbach and Weissbach (1988) Methods for Plant Molecular Biology, Academic Press, New York NY, pp 421-463.

An alternative expression system which could be used to express TUPRO is an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The TUPRO coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of TUPRO will render the polyhedrin gene inactive and

produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect S. frugiperda cells or Trichoplusia larvae in which TUPRO is expressed (Smith et al (1983) J Virol 46:584; Engelhard EK et al (1994) Proc Nat Acad Sci 91:3224-7).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a TUPRO coding sequence may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing TUPRO in infected host cells (Logan and Shenk (1984) Proc Natl Acad Sci 81:3655-59). In addition, transcription enhancers, such as the rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of a TUPRO sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where TUPRO, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf D et al (1994) Results Probl Cell Differ 20:125-62; Bittner et al (1987) Methods in Enzymol 153:516-544).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express TUPRO

may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M et al (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy I et al (1980) Cell 22:817-23) genes which can be employed in tk- or aprt- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler M et al (1980) Proc Natl Acad Sci 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin F et al (1981) J Mol Biol 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman SC and RC Mulligan (1988) Proc Natl Acad Sci 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, ß glucuronidase and its substrate, GUS, and luciferase and its substrate, luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes CA et al (1995) Methods Mol Biol 55:121-131).

Identification of Transformants Containing the Polynucleotide Sequence

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression should be confirmed. For example, if the TUPRO is inserted within a marker gene sequence, recombinant cells containing TUPRO can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a TUPRO sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem TUPRO as well.

Alternatively, host cells which contain the coding sequence for TUPRO and express TUPRO may be identified by a variety of procedures known to

those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridization and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of the nucleic acid or protein.

The presence of the polynucleotide sequence encoding TUPRO can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes, portions or fragments of polynucleotides encoding TUPRO. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the TUPRO-encoding sequence to detect transformants containing DNA or RNA encoding TUPRO. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of TUPRO, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TUPRO is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton R et al (1990, Serological Methods, a Laboratory Manual, APS Press, St Paul MN) and Maddox DE et al (1983, J Exp Med 158:1211).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding TUPRO include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the TUPRO sequence, or any portion of it, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labeled nucleotides.

A number of companies such as Pharmacia Biotech (Piscataway NJ), Promega (Madison WI), and US Biochemical Corp (Cleveland OH) supply commercial kits and protocols for these procedures. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. Patents teaching the use of such labels include US Patents 3,817,837; 3,850,752; 3,939,350; 3,996,345;

4,277,437; 4,275,149 and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No. 4,816,567 incorporated herein by reference.

Purification of TUPRO

Host cells transformed with a nucleotide sequence encoding TUPRO may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides encoding TUPRO can be designed with signal sequences which direct secretion of TUPRO through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may join TUPRO to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins (Kroll DJ et al (1993) DNA Cell Biol 12:441-53; cf discussion of vectors infra containing fusion proteins).

TUPRO may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor XA or enterokinase (Invitrogen, San Diego CA) between the purification domain and TUPRO is useful to facilitate purification. One such expression vector provides for expression of a fusion protein compromising a TUPRO and contains nucleic acid encoding 6 histidine residues followed by thioredoxin and an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromotography as described in Porath et al (1992) Protein Expression and Purification 3: 263-281) while the enterokinase cleavage site provides a means for purifying TUPRO from the fusion protein.

In addition to recombinant production, fragments of TUPRO may be produced by direct peptide synthesis using solid-phase techniques (cf Stewart et al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J (1963) J Am Chem Soc 85:2149-2154). In vitro protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City CA) in accordance with the instructions provided by the manufacturer. Various fragments of TUPRO

may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

Uses of TUPRO and Polynucleotides Encoding TUPRO

The rationale for use of the novel nucleotide and polypeptide sequences disclosed herein is based in part on the chemical and structural homology among TUPROA, human tumor protein D52 (GI 790225; Byrne et al, supra), and C. elegans F13E6.1 (GI 1072344; Wilson et al, supra) and between TUPROB and C. elegans ORF ZK418.5 (GI 470373; Wilson et al, supra). In addition, northern analysis disclosed herein indicates that TUPRO molecules are expressed in cells or tissue derived from many types of human cancer (Figs. 3 and 4).

Mutations in tumor genes are very often found in human tumors and in many instances are thought to be critical to both the initiation of tumor development and to the tumor's ability to survive chemotherapy. Tumor proteins may be essential for tumor development or may enable tumors to withstand chemotherapy. They are therefore potential targets for novel diagnostics and therapeutics. Accordingly, the novel tumor protein TUPRO or a TUPRO derivative, may be used to diagnose, prevent, or treat cancer. In conditions where TUPRO protein activity is not desirable, cells could be treated with an antagonist of TUPRO. Thus, TUPRO antagonists may be used to inactivate TUPRO-specific tumor processes.

TUPRO Antibodies

TUPRO-specific antibodies are useful for the diagnosis of conditions and diseases associated with expression of TUPRO. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library. Neutralizing antibodies, ie, those which inhibit dimer formation, are especially preferred for diagnostics and therapeutics.

TUPRO for antibody induction does not require biological activity; however, the protein fragment, or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids. Preferably, they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of TUPRO amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. Procedures well known in the art can be used for the production of antibodies to TUPRO.

For the production of antibodies, various hosts including goats,

rabbits, rats, mice, etc may be immunized by injection with TUPRO or any portion, fragment or oligopeptide which retains immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are potentially useful human adjuvants.

Monoclonal antibodies to TUPRO may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein (1975 Nature 256:495-497), the human B-cell hybridoma technique (Kosbor et al (1983) Immunol Today 4:72; Cote et al (1983) Proc Natl Acad Sci 80:2026-2030) and the EBV-hybridoma technique (Cole et al (1985) Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, New York NY, pp 77-96).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison et al (1984) Proc Natl Acad Sci 81:6851-6855; Neuberger et al (1984) Nature 312:604-608; Takeda et al (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies (US Patent No. 4,946,778) can be adapted to produce TUPRO-specific single chain antibodies

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi et al (1989, Proc Natl Acad Sci 86: 3833-3837), and Winter G and Milstein C (1991; Nature 349:293-299).

Antibody fragments which contain specific binding sites for TUPRO may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse WD et al (1989) Science 256:1275-1281).

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically

involve the formation of complexes between TUPRO and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific TUPRO protein is preferred, but a competitive binding assay may also be employed. These assays are described in Maddox DE et al (1983, J Exp Med 158:1211).

Diagnostic Assays Using TUPRO Specific Antibodies

Particular TUPRO antibodies are useful for the diagnosis of conditions or diseases characterized by expression of TUPRO or in assays to monitor patients being treated with TUPRO, agonists or inhibitors. Diagnostic assays for TUPRO include methods utilizing the antibody and a label to detect TUPRO in human body fluids or extracts of cells or tissues. The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, the polypeptides and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known, several of which were described above.

A variety of protocols for measuring TUPRO, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TUPRO is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp Med 158:1211).

In order to provide a basis for diagnosis, normal or standard values for TUPRO expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with antibody to TUPRO under conditions suitable for complex formation which are well known in the art. The amount of standard complex formation may be quantified by comparing various artificial membranes containing known quantities of TUPRO with both control and disease samples from biopsied tissues. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values establishes the presence of disease state.

Drug Screening

TUPRO, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic compounds in any of a variety

of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between TUPRO and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the TUPRO is described in detail in "Determination of Amino Acid Sequence Antigenicity" by Geysen HM, WO Application 84/03564, published on September 13, 1984, and incorporated herein by reference. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with fragments of TUPRO and washed. Bound TUPRO is then detected by methods well known in the art. Purified TUPRO can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding TUPRO specifically compete with a test compound for binding TUPRO. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TUPRO.

Diagnostic and Therapeutic Uses of the Polynucleotide

A polynucleotide encoding TUPRO, or any part thereof, may be used for diagnostic and/or therapeutic purposes. For diagnostic purposes, polynucleotides encoding TUPRO of this invention may be used to detect and quantitate gene expression in biopsied tissues in which expression of TUPRO may be implicated. The diagnostic assay is useful to distinguish between absence, presence, and excess expression of TUPRO and to monitor regulation of TUPRO levels during therapeutic intervention. Included in the scope of the invention are oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs.

Another aspect of the subject invention is to provide for hybridization or PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TUPRO or closely related molecules. The specificity of the probe, whether it is made from a highly specific region, eg, 10 unique nucleotides in the 5' regulatory region, or a less specific region, eg, especially in the 3' region, and the stringency of the hybridization or amplification (maximal, high, intermediate or low) will determine whether the probe identifies only naturally occurring sequences encoding TUPRO, alleles or related sequences.

Probes may also be used for the detection of related sequences and should preferably contain at least 50% of the nucleotides from any of these TUPRO encoding sequences. The hybridization probes of the subject invention may be derived from the nucleotide sequence of SEQ ID NO:2, SEQ ID NO:4, or from genomic sequence including promoter, enhancer elements and introns of the naturally occurring TUPRO. Hybridization probes may be labeled by a variety of reporter groups, including radionuclides such as 32P or 35S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Other means for producing specific hybridization probes for DNAs encoding TUPRO include the cloning of nucleic acid sequences encoding TUPRO or TUPRO derivatives into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides.

Polynucleotide sequences encoding TUPRO may be used for the diagnosis of conditions or diseases with which the expression of TUPRO is associated. For example, polynucleotide sequences encoding TUPRO may be used in hybridization or PCR assays of fluids or tissues from biopsies to detect TUPRO expression. The form of such qualitative or quantitative methods may include Southern or northern analysis, dot blot or other membrane-based technologies; PCR technologies; dip stick, pIN, chip and ELISA technologies. All of these techniques are well known in the art and are the basis of many commercially available diagnostic kits.

The nucleotide sequences encoding TUPRO disclosed herein provide the basis for assays that detect activation or induction associated with cancer. The nucleotide sequence encoding TUPRO may be labeled by methods known in the art and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After an incubation period, the sample is washed with a compatible fluid which optionally contains a dye (or other label requiring a developer) if the nucleotide has been labeled with an enzyme. After the compatible fluid is rinsed off, the dye is quantitated and compared with a standard. If the amount of dye in the biopsied or extracted sample is significantly elevated over that of a comparable control sample, the nucleotide sequence has hybridized with nucleotide sequences in the sample, and the presence of elevated levels of nucleotide sequences encoding TUPRO in the sample indicates the presence of the associated disease.

Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regime in animal studies, in clinical trials, or in

monitoring the treatment of an individual patient. In order to provide a basis for the diagnosis of disease, a normal or standard profile for TUPRO expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with TUPRO, or a portion thereof, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained for normal subjects with a dilution series of TUPRO run in the same experiment where a known amount of a substantially purified TUPRO is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients afflicted with TUPRO-associated diseases. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established, a therapeutic agent is administered and a treatment profile is generated. Such assays may be repeated on a regular basis to evaluate whether the values in the profile progress toward or return to the normal or standard pattern. Successive treatment profiles may be used to show the efficacy of treatment over a period of several days or several months.

PCR, as described in US Patent Nos. 4,683,195 and 4,965,188, provides additional uses for oligonucleotides based upon the TUPRO sequence. Such oligomers are generally chemically synthesized, but they may be generated enzymatically or produced from a recombinant source. Oligomers generally comprise two nucleotide sequences, one with sense orientation (5'->3') and one with antisense (3'<-5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Additionally, methods which may be used to quantitate the expression of a particular molecule include radiolabeling (Melby PC et al 1993 J Immunol Methods 159:235-44) or biotinylating (Duplaa C et al 1993 Anal Biochem 229-36) nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated. Quantitation of multiple samples may be speeded up by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation. For example, the presence of a relatively high amount of TUPRO in extracts of biopsied tissues may indicate the onset of cancer. A definitive diagnosis of this type may allow health professionals to begin aggressive treatment and prevent further worsening of the condition. Similarly, further assays can be used to monitor the progress of a patient

during treatment. Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known such as the triplet genetic code, specific base pair interactions, and the like.

Based upon its homology to the gene encoding D52, and its expression profile, polynucleotide sequences encoding TUPRO disclosed herein may be useful in the treatment of cancer.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense polynucleotides of the gene encoding TUPRO. See, for example, the techniques described in Sambrook et al (supra) and Ausubel et al (supra).

The polynucleotides comprising full length cDNA sequence and/or its regulatory elements enable researchers to use sequences encoding TUPRO as an investigative tool in sense (Youssoufian H and HF Lodish 1993 Mol Cell Biol 13:98-104) or antisense (Eguchi et al (1991) Annu Rev Biochem 60:631-652) regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers, or larger fragments, can be designed from various locations along the coding or control regions.

Genes encoding TUPRO can be turned off by transfecting a cell or tissue with expression vectors which express high levels of a desired TUPRO-encoding fragment. Such constructs can flood cells with untranslatable sense or antisense sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until all copies are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector (Mettler I, personal communication) and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA or PNA, to the control regions of gene encoding TUPRO, ie, the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, eg, between -10 and +10 regions of the leader sequence, are preferred. The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules.

Recent therapeutic advances using triplex DNA were reviewed by Gee JE et al (In: Huber BE and BI Carr (1994) Molecular and Immunologic Approaches, Futura Publishing Co, Mt Kisco NY).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding TUPRO.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding TUPRO. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine and wybutosine as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Methods for introducing vectors into cells or tissues include those methods discussed infra and which are equally suitable for in vivo, in vitro

and ex vivo therapy. For ex vivo therapy, vectors are introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient is presented in US Patent Nos. 5,399,493 and 5,437,994, disclosed herein by reference. Delivery by transfection and by liposome are quite well known in the art.

Furthermore, the nucleotide sequences for TUPRO disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including but not limited to such properties as the triplet genetic code and specific base pair interactions.

Detection and Mapping of Related Polynucleotide Sequences

The nucleic acid sequence for TUPRO can also be used to generate hybridization probes for mapping the naturally occurring genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include in situ hybridization to chromosomal spreads, flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price CM (1993; Blood Rev 7:127-34) and Trask BJ (1991; Trends Genet 7:149-54).

The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY. Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding TUPRO on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. For example, a sequence tagged site based map of the human genome was recently published by the Whitehead-MIT Center for Genomic Research (Hudson TJ et al (1995) Science 270:1945-1954). Often the placement of a gene on the chromosome of another mammalian species such as mouse (Whitehead Institute/MIT Center for Genome Research, Genetic Map of the Mouse, Database Release 10, April 28, 1995) may reveal associated

markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

Pharmaceutical Compositions

The present invention relates to pharmaceutical compositions which may comprise nucleotides, proteins, antibodies, agonists, antagonists, or inhibitors, alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. Any of these molecules can be administered to a patient alone, or in combination with other agents, drugs or hormones, in pharmaceutical compositions where it is mixed with excipient(s) or pharmaceutically acceptable carriers. In one embodiment of the present invention, the pharmaceutically acceptable carrier is pharmaceutically inert.

Administration of Pharmaceutical Compositions

Administration of pharmaceutical compositions is accomplished orally or parenterally. Methods of parenteral delivery include topical, intra-arterial (directly to the tumor), intramuscular, subcutaneous, intramedullary, intrathecal, intraventricular, intravenous, intraperitoneal, or intranasal administration. In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of "Remington's Pharmaceutical Sciences" (Maack Publishing Co, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical

compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; and gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores are provided with suitable coatings such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, ie, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders such as lactose or starches, lubricants such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations for parenteral administration include aqueous solutions of active compounds. For injection, the pharmaceutical compositions of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally,

the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Manufacture and Storage

The pharmaceutical compositions of the present invention may be manufactured in a manner that known in the art, eg, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder in 1mM-50 mM histidine, 0.1%-2% sucrose, 2%-7% mannitol at a pH range of 4.5 to 5.5 that is combined with buffer prior to use.

After pharmaceutical compositions comprising a compound of the invention formulated in a acceptable carrier have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of TUPRO, such labeling would include amount, frequency and method of administration.

Therapeutically Effective Dose

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, eg, of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model is also used to achieve a desirable concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or its antibodies, antagonists, or inhibitors which ameliorate the symptoms or condition. Therapeutic efficacy and toxicity of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, eg, ED50 (the dose therapeutically effective in 50% of

the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage is chosen by the individual physician in view of the patient to be treated. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Additional factors which may be taken into account include the severity of the disease state, eg, tumor size and location; age, weight and gender of the patient; diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long acting pharmaceutical compositions might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

It is contemplated, for example, that TUPRO or a TUPRO derivative can be delivered in a suitable formulation to stop the progression of cancer.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

EXAMPLES

I cDNA Library Construction

TLYMNOR01 cDNA Library

The TLYMNOR01 cDNA library was constructed using RNA isolated from non-adherent peripheral blood mononuclear cells obtained from a 24-year-old Caucasian male. The cells were purified on Ficoll Hypaque, then harvested,

lysed in GuSCN, and spun through CsCl to obtain RNA for library construction. The RNA was primed with oligo dT and cDNA was synthesized from the mRNA. Synthetic adaptor oligonucleotides were ligated onto cDNA ends enabling its insertion into Uni-ZAP™ XR vector system (Stratagene), allowing high efficiency unidirectional (sense orientation) lambda library construction and the convenience of a plasmid system with blue/white color selection to detect clones with cDNA insertions. Alternative unidirectional vectors include but are not limited to pcDNAI (Invitrogen) and pSHlox-1 (Novagen). Double-stranded cDNA was blunted, ligated to EcoRI adaptors, digested with XhoI, size-selected, and cloned into the XhoI and EcoRI sites of the Lambda UniZAP vector.

The cDNA library can be screened with either DNA probes or antibody probes and the pBluescript® phagemid (Stratagene) can be rapidly excised in vivo. The phagemid allows the use of a plasmid system for easy insert characterization, sequencing, site-directed mutagenesis, the creation of unidirectional deletions and expression of fusion proteins. The library phage particles were infected into E. coli host strain XL1-Blue® (Stratagene), which has a high transformation efficiency, increasing the probability of obtaining rare, under-represented clones in the cDNA library.

PROSTUT01 cDNA library

The PROSTUT01 cDNA library was constructed from prostate tumor tissue removed from a 50 year old Caucasian male (lot #0024A; Mayo Clinic, Rochester MN) by retropubic prostatectomy. The pathology report indicated Mayo grade 3 of 4 adenocarcinoma (Gleason grade 3+3). The tumor perforated and involved periprostatic tissue. There was also perineural invasion. The patient history revealed dysuria and treatment with an antibiotic. The patient had also reported a syncopal episode which did not require further treatment.

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron-PT 3000 (Brinkmann Instruments, Inc. Westbury NY) in guanidinium isothiocyanate solution. The lysate was extracted once with acid phenol at pH 4.0 per Stratagene's RNA isolation protocol (Stratagene Inc, San Diego CA). The lysate was re-extracted once more with phenol chloroform at pH 4.0. The RNA was then precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in DEPC-treated water and DNase treated for 25 min at 37°C. The RNA was again extracted three times with an equal volume of acid phenol, and reprecipitated using conditions described above. The mRNA was isolated using the Qiagen Oligotex kit (QIAGEN Inc, Chatsworth CA) and used to construct the cDNA library.

First strand cDNA synthesis was accomplished using an oligo d(T)

primer/linker which also contained an XhoI restriction site. Second strand synthesis was performed using a combination of DNA polymerase I, E. coli ligase and RNase H, followed by the addition of an EcoRI adaptor to the blunt ended cDNA. The EcoRI adapted, double-stranded cDNA was then digested with XhoI restriction enzyme and fractionated on Sephacryl S400 to obtain sequences which exceeded 400 bp in size. The size selected cDNAs were inserted into the LambdaZap® vector system (Stratagene, La Jolla CA); and the vector which contains the pBluescript™ phagemid (Stratagene) was transformed into cells of E. coli, strain XL1-BlueMRF™ (Stratagene).

The phagemid forms of individual cDNA clones were obtained by the in vivo excision process. Enzymes from both pBluescript and a cotransformed f1 helper phage nicked the DNA, initiated new DNA synthesis, and created the smaller, single-stranded circular phagemid DNA molecules which contained the cDNA insert. The phagemid DNA was released, purified, and used to reinfect fresh host cells (SOLR, Stratagene). Presence of the phagemid which carries the gene for ß-lactamase allowed transformed bacteria to grow on medium containing ampicillin.

II Isolation and Sequencing of cDNA Clones

The phagemid forms of individual cDNA clones were obtained by the in vivo excision process, in which the host bacterial strain was coinfected with both the lambda library phage and an f1 helper phage. Proteins derived from both the library-containing phage and the helper phage nicked the lambda DNA, initiated new DNA synthesis from defined sequences on the lambda target DNA and created a smaller, single stranded circular phagemid DNA molecule that included all DNA sequences of the pBluescript® plasmid and the cDNA insert. The phagemid DNA was secreted from the cells and purified, then used to re-infect fresh host cells, where the double stranded phagemid DNA was produced. Because the phagemid carries the gene for β-lactamase, the newly-transformed bacteria are selected on medium containing ampicillin.

Phagemid DNA was purified using the QIAwell-8 Plasmid Purification System™ from QIAGEN®, QIAwell PLUS™, or QIAwell ULTRA™ DNA Purification System (QIAGEN Inc, Chatsworth CA). The DNA was eluted from the purification resin already prepared for DNA sequencing and other analytical manipulations.

The cDNA inserts from random isolates of the library were sequenced in part by the method of Sanger F and AR Coulson (1975; J Mol Biol 94:441f), using a Hamilton Micro Lab 2200 (Hamilton, Reno NV) in combination with four Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown MA) and Applied Biosystems 377 or 373 DNA Sequencing Systems (Perkin Elmer), and the reading frame was determined.

III Homology Searching of cDNA Clones and Their Deduced Proteins

Each cDNA was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (TRW Inc, Los Angeles CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT- 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

BLAST, which stands for Basic Local Alignment Search Tool (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10), was used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. BLAST is useful for matches which do not contain gaps. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labelled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al. supra).

Analogous computer techniques using BLAST (Altschul SF 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte, Palo Alto CA). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

and it takes into acccount both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of the search are reported as a list of 1) libraries in which the full length sequence, or parts thereof, is represented 2) the abundance of the sequence, and 3) the percent abundance. Abundance directly reflects the number of times a particular transcript is present in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the library.

V Extension of TUPRO-Encoding Polynucleotides to Full Length or to Recover Regulatory Elements

Full length TUPRO-encoding nucleic acid sequences (SEQ ID NO:2 or SEQ ID NO:4) are used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers allow the extension of the known TUPRO-encoding sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest (US Patent Application 08/487,112, filed June 7, 1995, specifically incorporated by reference). The initial primers are designed from the cDNA using OLIGO® 4.06 Primer Analysis Software (National Biosciences), or another appropriate program, to

be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat step 4-6 for 15 additional cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat step 8-10 for 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were selected and cut out of the gel. Further purification involves using a commercial gel extraction method such as QIAQuick™ (QIAGEN Inc). After recovery of the DNA, Klenow enzyme was used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation, the products are redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook J et al, supra). After incubation for one hour at 37° C, the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook J et al, supra) containing 2xCarb. The following day, several colonies are randomly picked from each plate and

cultured in 150 μl of liquid LB/2xCarb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μl of each sample is transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer and one or both of the gene specific primers used for the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 or SEQ ID NO:4 are employed to screen cDNAs, genomic DNAs or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 mCi of [γ-^{32}P] adenosine triphosphate (Amersham, Chicago IL) and T4 polynucleotide kinase (DuPont NEN®, Boston MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 super fine resin column (Pharmacia). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester NY) is exposed to the blots in a Phosphoimager cassette (Molecular Dynamics,

Sunnyvale CA) for several hours, hybridization patterns are compared visually.

VII Antisense Molecules

The TUPRO-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring TUPRO. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequence of TUPRO, as shown in Figs. 1A, 1B, 2A, and 2B, is used to inhibit expression of naturally occurring TUPRO. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in Figures 1A, 1B, 2A, and 2B, and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of a TUPRO-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the leader and 5' sequence of SEQ ID NO:2 or SEQ ID NO:4, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or early coding sequence of the polypeptide as shown in Figures 1A, 1B, 2A and 2B.

VIII Expression of TUPRO

Expression of the TUPRO is accomplished by subcloning the cDNAs into appropriate vectors and transfecting the vectors into host cells. In this case, the cloning vector, pSport, previously used for the generation of the cDNA library is used to express TUPRO in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met and the subsequent 7 residues of ß-galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transfected bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first seven residues of ß-galactosidase, about 5 to 15 residues of linker, and the full length TUPRO-encoding sequence. The signal sequence directs the secretion of TUPRO into the bacterial growth media which can be used directly in the following assay for activity.

IX TUPRO Activity

TUPRO's ability to form either homodimers or heterodimers with human D52 can be measured by a common immunoprecipitation technique, such as described by Heymach JV et al (1995, J Biol Chem 270: 12297-12304). Human

D52 and D53 rabbit antisera is raised against peptides corresponding to internal sequences in which the two proteins share no amino acid identity. COS cells are transiently transfected with vector alone or expression plasmids for human D52, TUPRO, or both, and conditioned media from the transfectants is analyzed directly or after immunoprecipitation with anti-D52 or anti-TUPRO monoclonal antibody. Samples are run by SDS-PAGE and immunoblotted. Duplicate immunoblots are probed with either D52 or TUPRO antisera (Heymach et al, supra). Secondary antibodies conjugated to peroxidase are used to reveal homodimers and heterodimer formation.

X Production of TUPRO Specific Antibodies

TUPRO substantially purified using PAGE electrophoresis (Sambrook, supra) is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence translated from TUPRO is analyzed using DNAStar software (DNAStar Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Analysis to select appropriate epitopes, such as those near the C-terminus or in hydrophilic regions (shown in Figs. 7 and 9) is described by Ausubel FM et al (supra).

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel FM et al, supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring TUPRO Using Specific Antibodies

Naturally occurring or recombinant TUPRO is substantially purified by immunoaffinity chromatography using antibodies specific for TUPRO. An immunoaffinity column is constructed by covalently coupling TUPRO antibody to an activated chromatographic resin such as CnBr-activated Sepharose (Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TUPRO is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TUPRO (eg, high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TUPRO binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope such as urea

or thiocyanate ion), and TUPRO is collected.

XII Identification of Molecules Which Interact with TUPRO

TUPRO, or biologically active fragments thereof, are labelled with ^{125}I Bolton-Hunter reagent (Bolton, AE and Hunter, WM (1973) Biochem J 133: 529). Candidate molecules previously arrayed in the wells of a 96 well plate are incubated with the labelled TUPRO, washed and any wells with labelled TUPRO complex are assayed. Data obtained using different concentrations of TUPRO are used to calculate values for the number, affinity, and association of TUPRO with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Bandman, Olga
Au-Young, Janice
Goli, Surya K.
Hillman, Jennifer.
Zweiger, Gary B.

(ii) TITLE OF THE INVENTION: A NOVEL TUMOR PROTEIN

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: U.S.
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0126 US

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415-855-0555
(B) TELEFAX: 415-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 204 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Glu Ala Gln Ala Gln Gly Leu Leu Glu Thr Glu Pro Leu Gln Gly
 1               5                  10                  15
Thr Asp Glu Asp Ala Val Ala Ser Ala Asp Phe Ser Ser Met Leu Ser
            20                  25                  30
Glu Glu Glu Lys Glu Glu Leu Lys Ala Glu Leu Val Gln Leu Glu Asp
        35                  40                  45
Glu Ile Thr Thr Leu Arg Gln Val Leu Ser Ala Lys Glu Arg His Leu
    50                  55                  60
Val Glu Ile Lys Gln Lys Leu Gly Met Asn Leu Met Asn Glu Leu Lys
65                  70                  75                  80
Gln Asn Phe Ser Lys Ser Trp His Asp Met Gln Thr Thr Thr Ala Tyr
                85                  90                  95
Lys Lys Thr His Glu Thr Leu Ser His Ala Gly Gln Lys Ala Thr Ala
            100                 105                 110
Ala Phe Ser Asn Val Gly Thr Ala Ile Ser Lys Lys Phe Gly Asp Met
        115                 120                 125
Ser Tyr Ser Ile Arg His Ser Ile Ser Met Pro Ala Met Arg Asn Ser
    130                 135                 140
Pro Thr Phe Lys Ser Phe Glu Glu Arg Val Glu Thr Thr Val Thr Ser
145                 150                 155                 160
Leu Lys Thr Lys Val Gly Gly Thr Asn Pro Asn Gly Gly Ser Phe Glu
                165                 170                 175
Glu Val Leu Ser Ser Thr Ala His Ala Ser Ala Gln Ser Leu Ala Gly
            180                 185                 190
Gly Ser Arg Arg Thr Lys Glu Glu Glu Leu Gln Cys
        195                 200
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 790 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GGGCCAGCTG CGTTCTGAGC CTGGGCGCAG CTACCATCTG CTCTGGGAAG CACCAGGGTG     60
TCCCCGCCGC CCTCAGCTCG AAGTCAGCCA CCATGGAGGC GCAGGCACAA GGTTTGTTGG    120
AGACTGAACC GTTGCAAGGA ACAGACGAAG ATGCAGTAGC CAGTGCTGAC TTCTCTAGCA    180
TGCTCTCTGA GGAGGAAAAG GAAGAGTTAA AAGCAGAGTT AGTTCAGCTA GAAGACGAAA    240
TTACAACACT ACGACAAGTT TTGTCAGCGA AAGAAAGGCA TCTAGTTGAG ATAAAACAAA    300
AACTCGGCAT GAACCTGATG AATGAATTAA AACAGAACTT CAGCAAAAGC TGGCATGACA    360
TGCAGACTAC CACTGCCTAC AAGAAAACAC ATGAAACCCT GAGTCACGCA GGGCAAAAGG    420
CAACTGCAGC TTTCAGCAAC GTTGGAACGG CCATCAGCAA GAAGTTCGGA GACATGAGTT    480
ACTCCATTCG CCATTCCATA AGTATGCCTG CTATGAGGAA TTCTCCTACT TTCAAATCAT    540
TTGAGGAGAG GGTTGAGACA ACTGTCACAA GCCTCAAGAC GAAAGTAGGC GGTACGAACC    600
CTAATGGAGG CAGTTTTGAG GAGGTCCTCA GCTCCACGGC CCATGCCAGT GCCCAGAGCT    660
TGGCAGGAGG CTCCCGGCGG ACCAAGGAGG AGGAGCTGCA GTGCTAAGTC CAGCCAGCGT    720
```

```
GCAGTGCATC CAGAAACCGG CCACTACCCA GCCCATCTNT GCCTGTGCTT ATCCAGATAA    780
GAAGACCAAA                                                           790
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 245 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
 (A) LIBRARY:
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Thr Leu Phe His Phe Gly Asn Cys Phe Ala Leu Ala Tyr Phe Pro
 1               5                  10                  15
Tyr Phe Ile Thr Tyr Lys Cys Ser Gly Leu Ser Glu Tyr Asn Ala Phe
            20                  25                  30
Trp Lys Cys Val Gln Ala Gly Val Thr Tyr Leu Phe Val Gln Leu Cys
        35                  40                  45
Lys Met Leu Phe Leu Ala Thr Phe Phe Pro Thr Trp Glu Gly Gly Ile
    50                  55                  60
Tyr Asp Phe Ile Gly Glu Phe Met Lys Ala Ser Val Asp Val Ala Asp
65                  70                  75                  80
Leu Ile Gly Leu Asn Leu Val Met Ser Arg Asn Ala Gly Lys Gly Glu
                85                  90                  95
Tyr Lys Ile Met Val Ala Ala Leu Gly Trp Ala Thr Ala Glu Leu Ile
            100                 105                 110
Met Ser Arg Cys Ile Pro Leu Trp Val Gly Ala Arg Gly Ile Glu Phe
        115                 120                 125
Asp Trp Lys Tyr Ile Gln Met Ser Ile Asp Ser Asn Ile Ser Leu Val
    130                 135                 140
His Tyr Ile Val Ala Ser Ala Gln Val Trp Met Ile Thr Arg Tyr Asp
145                 150                 155                 160
Leu Tyr His Asn Phe Arg Pro Ala Val Leu Leu Leu Met Phe Leu Ser
                165                 170                 175
Val Tyr Lys Ala Phe Val Met Glu Thr Phe Val His Leu Cys Ser Leu
            180                 185                 190
Gly Ser Trp Ala Arg Leu Asp Ala Arg Ala Val Val Thr Gly Leu Leu
        195                 200                 205
Ala Leu Lys His Phe Gly Pro Val Cys Arg Arg Cys Gln Cys Ala Leu
    210                 215                 220
Leu Gly Leu Val Ser Gln Thr Leu Met Tyr Leu Phe Pro Ala Ser Leu
225                 230                 235                 240
Gln Val Leu Val Lys
                245
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 888 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
GTMGMKCGCG GGCCCCCGCC AGTCAGGTGG GTGCCAGGCC CTGGCCGTGG CGAAAGAGCC    60
GGCGGAGGGA GGACCCGCTC CCGGAGACGC CGCCTCGCGA TCCCCGCGCG GGCGGGACCG   120
GGCGGCCGGC ATCATGACCC TGTTTCACTT CGGGAACTGC TTCGCTCTTG CCTACTTCCC   180
CTACTTCATC ACCTACAAGT GCAGCGGCCT GTCCGAGTAC AACGCCTTCT GGAAATGCGT   240
CCAGGCTGGA GTCACCTACC TCTTTGTCCA ACTCTGCAAG ATGCTGTTCT TGGCCACTTT   300
CTTTCCCACC TGGGAAGGCG GCATCTATGA CTTCATTGGG GAGTTCATGA AGGCCAGCGT   360
GGATGTGGCA GACCTGATAG GTCTAAACCT TGTCATGTCC CGGAATGCCG GCAAGGGAGA   420
GTACAAGATC ATGGTTGCTG CCCTGGGCTG GGCCACTGCT GAGCTTATTA TGTCCCGCTG   480
CATTCCCCTA TGGGTCGGAG CCCGGGGCAT TGAGTTTGAC TGGAAGTACA TCCAGATGAG   540
CATAGACTCC AACATCAGTC TGGTCCATTA CATCGTCGCG TCTGCTCAGG TCTGGATGAT   600
AACACGCTAT GATCTGTACC ACAACTTCCG GCCAGCTGTC CTTCTGCTGA TGTTCCTCAG   660
TGTCTACAAG GCCTTTGTTA TGGAGACCTT CGTCCACCTC TGCTCGCTGG GCAGTTGGGC   720
ARCTCTAMTG GCCCGAGCAG TGGTAACGGG GCTGCTGGCC CTCAAGCACT TTGGSCCTGT   780
ATGTCGSCGT TGTCAATGTG CACTYCTAGG CTTGGTGTCT CAGACATTGA TGTACCTTTT   840
CCCTGCCTCA CTCCAGGTTT TAGTGAAGTA AACAGTATTT GGAAAGTT                888
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 184 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 790225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Asp Arg Gly Glu Gln Gly Leu Leu Arg Thr Asp Pro Val Pro Glu
 1               5                  10                  15
Glu Gly Glu Asp Val Ala Ala Thr Ile Ser Ala Thr Glu Thr Leu Ser
            20                  25                  30
Glu Glu Glu Gln Glu Glu Leu Arg Arg Glu Leu Ala Lys Val Glu Glu
        35                  40                  45
Glu Ile Gln Thr Leu Ser Gln Val Leu Ala Ala Lys Glu Lys His Leu
    50                  55                  60
Ala Glu Ile Lys Arg Lys Leu Gly Ile Asn Ser Leu Gln Glu Leu Lys
65                  70                  75                  80
Gln Asn Ile Ala Lys Gly Trp Gln Asp Val Thr Ala Thr Ser Ala Tyr
                85                  90                  95
Lys Lys Thr Ser Glu Thr Leu Ser Gln Ala Gly Gln Lys Ala Ser Ala
            100                 105                 110
```

```
Ala Phe Ser Ser Val Gly Ser Val Ile Thr Lys Lys Leu Glu Asp Val
        115                 120                 125
Lys Asn Ser Pro Thr Phe Lys Ser Phe Glu Glu Lys Val Glu Asn Leu
    130                 135                 140
Lys Ser Lys Val Gly Gly Thr Lys Pro Ala Gly Gly Asp Phe Gly Glu
145                 150                 155                 160
Val Leu Asn Ser Ala Ala Asn Ala Ser Ala Thr Thr Thr Glu Pro Leu
                165                 170                 175
Pro Glu Lys Thr Gln Glu Ser Leu
            180
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 257 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 1072344

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Pro Lys Gly Asn Lys Lys Pro Asn Glu Lys Lys Glu Glu Leu Glu
 1               5                  10                  15
Lys Phe Ala Lys Glu Leu Gln Gly Ser Asp Ser Asp Glu Asp Ala Val
            20                  25                  30
Val Ile Glu Gln Pro Thr Val Glu Pro Lys Leu Pro Gln Asn Asp Ser
        35                  40                  45
Ser Ser Ser Asn Lys Ile Val Leu Ser Gln Ala Glu Lys Asp Leu Leu
    50                  55                  60
Arg Thr Glu Leu Asp Lys Thr Glu Glu Glu Ile Ser Thr Leu Lys Gln
65                  70                  75                  80
Val Leu Ser Ala Arg Gln Lys His Ala Ala Glu Leu Lys Arg Lys Leu
                85                  90                  95
Gly Leu Thr Pro Phe Ser Glu Leu Ser Gln Asp Ile Asn Arg Ser Leu
            100                 105                 110
Lys Thr Val Thr Asp Thr Asp Ala Cys Thr His Phe Ile Glu Ile Asn
        115                 120                 125
Ile Gln Lys Lys Lys Lys Gln Ser Met Tyr Tyr Ile Lys Arg Leu Ser
    130                 135                 140
Lys Asn Ile Gln Thr Val Pro Ile Leu Thr Ser Glu Lys Lys Arg Ile
145                 150                 155                 160
Leu His Ala Phe Ile Val Leu Lys Lys Lys Ser Ser Ile Leu Lys Ser
                165                 170                 175
Leu Leu Leu Trp Gln Gln Tyr Gln Lys Thr Ala Glu Val Ala Ala Ala
            180                 185                 190
Thr Ser Asp Thr Val Lys Glu Lys Trp Asn Asp Met Arg Asn Ser Ser
        195                 200                 205
Leu Phe Lys Ser Phe Glu Ser Lys Leu Gly Ser Ala Leu Asn Asn Ala
    210                 215                 220
Lys Met Ala Ala Ser Thr Ser Ile Asp His Leu Ala Gly Ala Ala Arg
225                 230                 235                 240
```

```
Gly Pro Ser Gln Thr Gly Thr Pro Val Ala Glu Glu Ala Lys Pro Ile
                245                 250                 255
Ser
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 470373

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Ser Phe Phe His Phe Ile Asn Cys Phe Ala Leu Ala Phe Ala Pro
 1               5                  10                  15
Tyr Phe Ile Val Tyr Lys Tyr Ser Gly Ile Asn Glu Tyr Ser Ser Ile
            20                  25                  30
Trp Lys Cys Ala Thr Ala Ser Gly Gly Tyr Leu Leu Thr Gln Leu Ala
        35                  40                  45
Lys Leu Leu Ile Ile Ala Thr Phe Phe Pro Ala Leu Asp Ser Glu Gly
    50                  55                  60
Phe Ser Ile Val Pro Glu Phe Leu Lys Ser Ser Ala Asp Ile Ile Asp
65                  70                  75                  80
Val Ile Gly Leu His Leu Leu Met Thr Asn Phe Leu Ala Gly Lys Gly
                85                  90                  95
Glu Val Arg Phe Val Val Gly Gly Leu Gly Trp Gly Phe Ala His Ser
            100                 105                 110
Val Ala His Arg Leu Val Leu Leu Trp Val Gly Ala Arg Gly Thr Ala
        115                 120                 125
Phe Thr Trp Arg Trp Val Gln Thr Ser Leu Asp Ser Ser Ala Asp Leu
    130                 135                 140
Leu Val Ile Val Ser Leu Ala Cys Leu Thr Trp Met Ile Thr Arg Thr
145                 150                 155                 160
Pro Asn Lys Phe Leu Val Ser Pro Ile Leu Ala Ile Thr Val Gln His
                165                 170                 175
Thr Phe Ser Leu Tyr Gly Trp Ser Leu Leu Ala Phe Arg Phe Ala Tyr
            180                 185                 190
Ser Ile Ala Thr Ala Ile Leu Thr Val Val Val Tyr Ser Ala Asn Arg
        195                 200                 205
Thr Ala Ser Thr Arg Lys Asn Glu
    210                 215
```

CLAIMS

1. A substantially purified human tumor protein comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.

2. An isolated and purified polynucleotide sequence encoding the protein of claim 1.

3. An isolated and purified polynucleotide sequence of claim 2 consisting of SEQ ID NO:2 or variants thereof.

4. A polynucleotide sequence which is complementary to SEQ ID NO:2 or variants thereof.

5. A recombinant expression vector containing the polynucleotide sequence of claim 2.

6. A recombinant host cell containing the vector of claim 5.

7. A method for producing a polypeptide comprising a polypeptide of SEQ ID NO:1, the method comprising the steps of:

a) culturing the host cell of claim 6 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

8. A pharmaceutical composition comprising a substantially purified human tumor protein having an amino acid sequence of SEQ ID NO:1 in conjunction with a suitable pharmaceutical carrier.

9. A purified antibody which binds specifically to the polypeptide of claim 1.

10. A purified antagonist which specifically regulates or modulates the activity of the polypeptide of claim 1.

11. A pharmaceutical composition comprising the antagonist of claim 10 in conjunction with a suitable pharmaceutical carrier.

12. A method for treating cancer comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.

13. A substantially purified human tumor protein comprising the amino acid sequence of SEQ ID NO:3 or fragments thereof.

14. An isolated and purified polynucleotide sequence encoding the protein of claim 13.

15. An isolated and purified polynucleotide sequence of claim 14 consisting of SEQ ID NO:4 or variants thereof.

16. A polynucleotide sequence which is complementary to SEQ ID NO:4 or variants thereof.

17. A recombinant expression vector containing the polynucleotide sequence of claim 14.

18. A recombinant host cell containing the vector of claim 17.

19. A method for producing a polypeptide comprising a polypeptide of SEQ ID NO:3, the method comprising the steps of:

a) culturing the host cell of claim 18 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

20. A pharmaceutical composition comprising a substantially purified human tumor protein having an amino acid sequence of SEQ ID NO:3 in conjunction with a suitable pharmaceutical carrier.

21. A purified antibody which binds specifically to the polypeptide of claim 13.

22. A purified antagonist which specifically regulates or modulates the activity of the polypeptide of claim 13.

23. A pharmaceutical composition comprising the antagonist of claim 22 in conjunction with a suitable pharmaceutical carrier.

24. A method for treating cancer comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 23.

ABSTRACT

NOVEL HUMAN TUMOR PROTEINS

The present invention provides novel human tumor proteins (collectively called TUPRO) and polynucleotides which identify and encode TUPRO. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding TUPRO. The invention also provides pharmaceutical compositions containing TUPRO or antagonists to TUPRO, and in the use of these compositions for the treatment of diseases associated with the expression of TUPRO. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding TUPRO for the treatment of diseases associated with the expression of TUPRO. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, to hybridize to the genomic sequence or transcripts of polynucleotides encoding TUPRO or anti-TUPRO antibodies which specifically bind to TUPRO.

```
            11          20          29          38          47          56
5' GCC AGC TGC GTT CTG AGC CTG GGC GCA GCT ACC ATC TGC TCT GGG AAG CAC CAG

            65          74          83          92         101         110
GGT GTC CCC GCC GCC CTC AGC TCG AAG TCA GCC ACC ATG GAG GCG CAG GCA CAA
                                                M   E   A   Q   A   Q

           119         128         137         146         155         164
GGT TTG TTG GAG ACT GAA CCG TTG CAA GGA ACA GAC GAA GAT GCA GTA GCC AGT
G   L   L   E   T   E   P   L   Q   G   T   D   E   D   A   V   A   S

           173         182         191         200         209         218
GCT GAC TTC TCT AGC ATG CTC TCT GAG GAG GAA AAG GAA GAG TTA AAA GCA GAG
A   D   F   S   S   M   L   S   E   E   E   K   E   E   L   K   A   E

           227         236         245         254         263         272
TTA GTT CAG CTA GAA GAC GAA ATT ACA ACA CTA CGA CAA GTT TTG TCA GCG AAA
L   V   Q   L   E   D   E   I   T   T   L   R   Q   V   L   S   A   K

           281         290         299         308         317         326
GAA AGG CAT CTA GTT GAG ATA AAA CAA AAA CTC GGC ATG AAC CTG ATG AAT GAA
E   R   H   L   V   E   I   K   Q   K   L   G   M   N   L   M   N   E

           335         344         353         362         371         380
TTA AAA CAG AAC TTC AGC AAA AGC TGG CAT GAC ATG CAG ACT ACC ACT GCC TAC
L   K   Q   N   F   S   K   S   W   H   D   M   Q   T   T   T   A   Y

           389         398         407         416         425         434
AAG AAA ACA CAT GAA ACC CTG AGT CAC GCA GGG CAA AAG GCA ACT GCA GCT TTC
K   K   T   H   E   T   L   S   H   A   G   Q   K   A   T   A   A   F

           443         452         461         470         479         488
AGC AAC GTT GGA ACG GCC ATC AGC AAG AAG TTC GGA GAC ATG AGT TAC TCC ATT
S   N   V   G   T   A   I   S   K   K   F   G   D   M   S   Y   S   I

           497         506         515         524         533         542
CGC CAT TCC ATA AGT ATG CCT GCT ATG AGG AAT TCT CCT ACT TTC AAA TCA TTT
R   H   S   I   S   M   P   A   M   R   N   S   P   T   F   K   S   F

           551         560         569         578         587         596
GAG GAG AGG GTT GAG ACA ACT GTC ACA AGC CTC AAG ACG AAA GTA GGC GGT ACG
E   E   R   V   E   T   T   V   T   S   L   K   T   K   V   G   G   T

           605         614         623         632         641         650
AAC CCT AAT GGA GGC AGT TTT GAG GAG GTC CTC AGC TCC ACG GCC CAT GCC AGT
N   P   N   G   G   S   F   E   E   V   L   S   S   T   A   H   A   S

           659         668         677         686         695         704
GCC CAG AGC TTG GCA GGA GGC TCC CGG CGG ACC AAG GAG GAG GAG CTG CAG TGC
A   Q   S   L   A   G   G   S   R   R   T   K   E   E   E   L   Q   C
```

FIGURE 1A

```
        713         722         731         740         749         758
TAA GTC CAG CCA GCG TGC AGT GCA TCC AGA AAC CGG CCA CTA CCC AGC CCA TCT

        767         776         785
NTG CCT GTG CTT ATC CAG ATA AGA AGA CCA AA 3'
```

FIGURE 1B

```
                10          19          28          37          46          55
5' TMG MKC GCG GGC CCC CGC CAG TCA GGT GGG TGC CAG GCC CTG GCC GTG GCG AAA

                64          73          82          91         100         109
   GAG CCG GCG GAG GGA GGA CCC GCT CCC GGA GAC GCC GCC TCG CGA TCC CCG CGC

                118         127         136         145         154         163
   GGG CGG GAC CGG GCG GCC GGC ATC ATG ACC CTG TTT CAC TTC GGG AAC TGC TTC
                                    M   T   L   F   H   F   G   N   C   F

                172         181         190         199         208         217
   GCT CTT GCC TAC TTC CCC TAC TTC ATC ACC TAC AAG TGC AGC GGC CTG TCC GAG
   A   L   A   Y   F   P   Y   F   I   T   Y   K   C   S   G   L   S   E

                226         235         244         253         262         271
   TAC AAC GCC TTC TGG AAA TGC GTC CAG GCT GGA GTC ACC TAC CTC TTT GTC CAA
   Y   N   A   F   W   K   C   V   Q   A   G   V   T   Y   L   F   V   Q

                280         289         298         307         316         325
   CTC TGC AAG ATG CTG TTC TTG GCC ACT TTC TTT CCC ACC TGG GAA GGC GGC ATC
   L   C   K   M   L   F   L   A   T   F   F   P   T   W   E   G   G   I

                334         343         352         361         370         379
   TAT GAC TTC ATT GGG GAG TTC ATG AAG GCC AGC GTG GAT GTG GCA GAC CTG ATA
   Y   D   F   I   G   E   F   M   K   A   S   V   D   V   A   D   L   I

                388         397         406         415         424         433
   GGT CTA AAC CTT GTC ATG TCC CGG AAT GCC GGC AAG GGA GAG TAC AAG ATC ATG
   G   L   N   L   V   M   S   R   N   A   G   K   G   E   Y   K   I   M

                442         451         460         469         478         487
   GTT GCT GCC CTG GGC TGG GCC ACT GCT GAG CTT ATT ATG TCC CGC TGC ATT CCC
   V   A   A   L   G   W   A   T   A   E   L   I   M   S   R   C   I   P

                496         505         514         523         532         541
   CTA TGG GTC GGA GCC CGG GGC ATT GAG TTT GAC TGG AAG TAC ATC CAG ATG AGC
   L   W   V   G   A   R   G   I   E   F   D   W   K   Y   I   Q   M   S

                550         559         568         577         586         595
   ATA GAC TCC AAC ATC AGT CTG GTC CAT TAC ATC GTC GCG TCT GCT CAG GTC TGG
   I   D   S   N   I   S   L   V   H   Y   I   V   A   S   A   Q   V   W

                604         613         622         631         640         649
   ATG ATA ACA CGC TAT GAT CTG TAC CAC AAC TTC CGG CCA GCT GTC CTT CTG CTG
   M   I   T   R   Y   D   L   Y   H   N   F   R   P   A   V   L   L   L

                658         667         676         685         694         703
   ATG TTC CTC AGT GTC TAC AAG GCC TTT GTT ATG GAG ACC TTC GTC CAC CTC TGC
   M   F   L   S   V   Y   K   A   F   V   M   E   T   F   V   H   L   C

                712         721         730         739         748         757
   TCG CTG GGC AGT TGG GCA RCT CTA MTG GCC CGA GCA GTG GTA ACG GGG CTG CTG
   S   L   G   S   W   A   X   L   X   A   R   A   V   V   T   G   L   L
```

FIGURE 2A

```
          766           775           784           793           802           811
GCC CTC AAG CAC TTT GGS CCT GTA TGT CGS CGT TGT CAA TGT GCA CTY CTA GGC
A   L   K   H   F   G   P   V   C   R   R   C   Q   C   A   L   L   G

          820           829           838           847           856           865
TTG GTG TCT CAG ACA TTG ATG TAC CTT TTC CCT GCC TCA CTC CAG GTT TTA GTG
L   V   S   Q   T   L   M   Y   L   F   P   A   S   L   Q   V   L   V

          874           883
AAG TAA ACA GTA TTT GGA AAG TT 3'
K
```

FIGURE 2B

Library	Lib Description	Abun	Pct Abun
BRAINOM02	brain, 55 M, NORM, WM	1	0.0454
UTRSNOT01	uterus, 59 F	1	0.0393
TLYMNOR01	lymphocytes (non-adher PBMNC), 24 M, RP	1	0.0372
BRSTNOT02	breast, 55 F, match to BRSTTUT01	2	0.0317
PROSTUT08	prostate tumor, 60 M, match to PROSNOT14	1	0.0266
PROSNOT14	prostate, 60 M, match to PROSTUT08	1	0.0256
LIVRNOM01	liver, 49 M, WM	1	0.0254
PROSNOT15	prostate, 66 M, match to PROSTUT10	1	0.0241
NERVMSM01	multiple sclerosis, 46 M, NORM, WM	1	0.0228
HNT2AGT01	hNT2 cell line, post-mitotic neurons	1	0.0190
BRAITUT02	brain tumor, metastasis, 58 M	1	0.0169
LIVSFEM02	liver/spleen, fetal M, NORM, WM	2	0.0053

Electronic Northern Results returned a total of 12 row(s).

FIGURE 3

Library	Lib Description	Abun	Pct Abun
COLNNOT22	colon, 56 F	2	0.0554
COLNPOT01	colon polyp, 40 F	2	0.0513
PROSNOT18	prostate, hyperplasia, 58 M	2	0.0513
MUSCNOT02	muscle, psoas, 12 M	1	0.0382
STOMTUT01	stomach tumor, 52 M, match to STOMNOT02	1	0.0368
SINTNOT02	small intestine, 55 F	1	0.0345
LVENNOT03	heart, left ventricle, 31 M	1	0.0337
MMLR3DT01	macrophages (adher PBMNC), M/F, 72-hr MLR	1	0.0331
PROSTUT01	prostate tumor, 50 M, match to PROSNOT02	1	0.0310
LUNGTUT03	lung tumor, 69 M, match to LUNGNOT15	1	0.0308
BLADTUT02	bladder tumor, 80 F, match to BLADNOT03	1	0.0305
BRAITUT08	brain tumor, astrocytoma, 47 M	2	0.0293
PROSTUT12	prostate tumor, 65 M, match to PROSNOT20	1	0.0279
BLADNOT04	bladder, 28 M	1	0.0278
TESTTUT02	testicular tumor, 31 M	1	0.0278
THYRNOT03	thyroid tumor, adenoma, 28 F	2	0.0277
SINTNOT13	small intestine, ileum, ulcerative cholitis, 25 F	1	0.0275
COLNTUT03	colon tumor, 62 M, match to COLNNOT16	1	0.0272
BLADTUT05	bladder tumor, 66 M, match to BLADNOT06	1	0.0268
KIDNTUT01	kidney tumor, Wilms, 8m F	1	0.0267
PENITUT01	penis tumor, carcinoma, 64 M	1	0.0267
COLNNOT23	colon, 16 M	1	0.0264
BRAITUT13	brain tumor, meningioma, 68 M	1	0.0262
LIVRTUT01	liver tumor, metastasis, 51 F	1	0.0259
PROSNOT14	prostate, 60 M, match to PROSTUT08	1	0.0256
BRSTTUT08	breast tumor, 45 F, match to BRSTNOT09	1	0.0254
BMARNOT03	bone marrow, 16 M	1	0.0242
RATRNOT02	heart, right atrium, 39 M	1	0.0236
PANCNOT01	pancreas, 29 M	1	0.0214
LUNGNOT04	lung, 2 M	1	0.0183
SYNORAT04	synovium, wrist, rheumatoid, 62 F	1	0.0174
PLACNOT02	placenta, fetal F	1	0.0168
BRSTNOT03	breast, 54 F, match to BRSTTUT02	1	0.0147
SPLNNOT04	spleen, 2 M	1	0.0128
PROSNOT06	prostate, 57 M, match to PROSTUT04	1	0.0114
LUNGFET03	lung, fetal F	1	0.0091

Electronic Northern Results returned a total of 36 row(s).

FIGURE 4

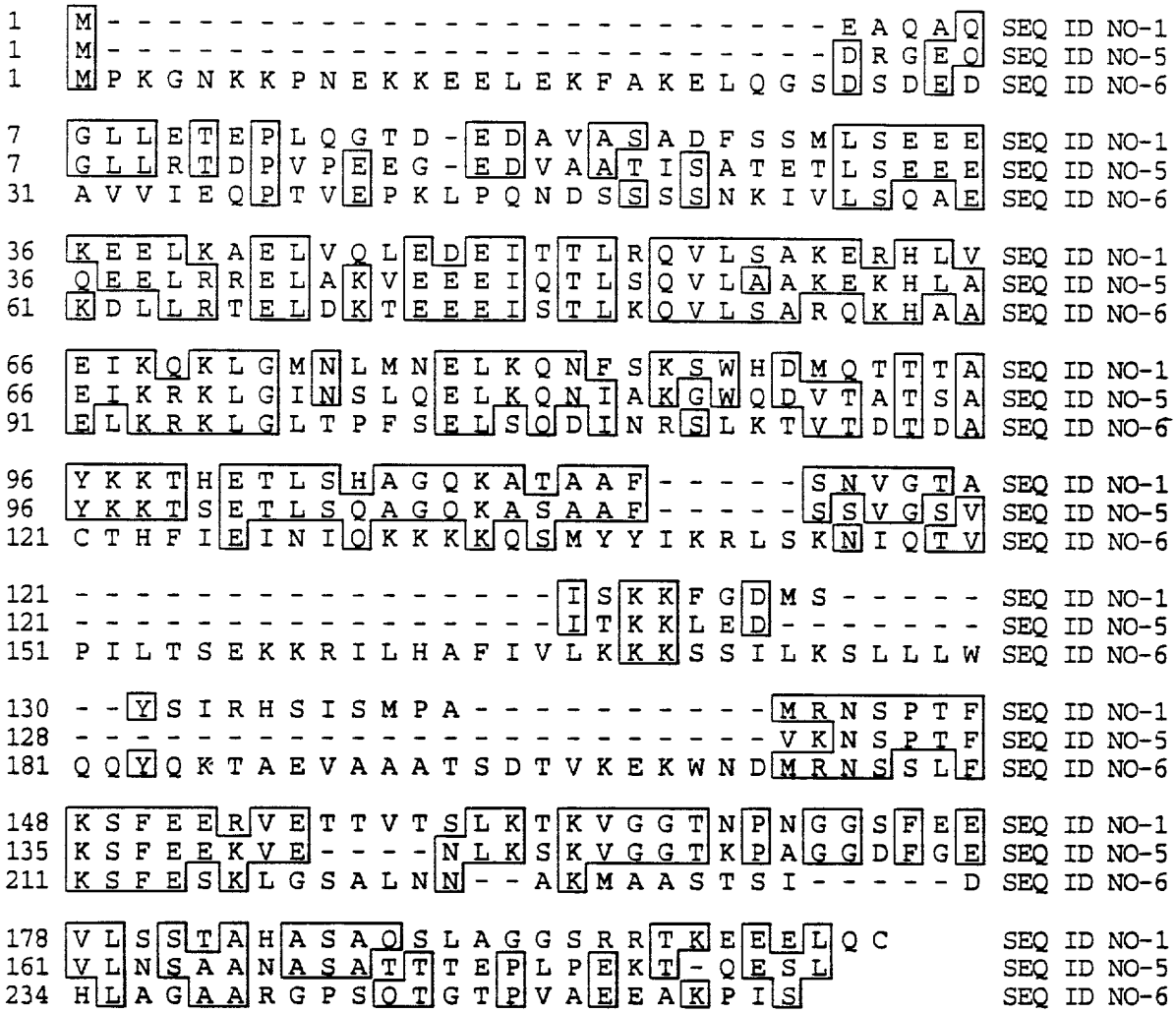
1 M - E A Q A Q SEQ ID NO-1
1 M - D R G E Q SEQ ID NO-5
1 M P K G N K K P N E K K E E L E K F A K E L Q G S D S D E D SEQ ID NO-6
7 G L L E T E P L Q G T D - E D A V A S A D F S S M L S E E E SEQ ID NO-1
7 G L L R T D P V P E E G - E D V A A T I S A T E T L S E E E SEQ ID NO-5
31 A V V I E Q P T V E P K L P Q N D S S S S N K I V L S Q A E SEQ ID NO-6
36 K E E L K A E L V Q L E D E I T T L R Q V L S A K E R H L V SEQ ID NO-1
36 Q E E L R R E L A K V E E E I Q T L S Q V L A A K E K H L A SEQ ID NO-5
61 K D L L R T E L D K T E E E I S T L K Q V L S A R Q K H A A SEQ ID NO-6
66 E I K Q K L G M N L M N E L K Q N F S K S W H D M Q T T T A SEQ ID NO-1
66 E I K R K L G I N S L Q E L K Q N I A K G W Q D V T A T S A SEQ ID NO-5
91 E L K R K L G L T P F S E L S Q D I N R S L K T V T D T D A SEQ ID NO-6
96 Y K K T H E T L S H A G Q K A T A A F - - - - - S N V G T A SEQ ID NO-1
96 Y K K T S E T L S Q A G Q K A S A A F - - - - - S S V G S V SEQ ID NO-5
121 C T H F I E I N I Q K K K K Q S M Y Y I K R L S K N I Q T V SEQ ID NO-6
121 - - - - - - - - - - - - - - - - - I S K K F G D M S - - - - - SEQ ID NO-1
121 - - - - - - - - - - - - - - - - - I T K K L E D - - - - - - - SEQ ID NO-5
151 P I L T S E K K R I L H A F I V L K K K S S I L K S L L L W SEQ ID NO-6
130 - - Y S I R H S I S M P A - - - - - - - - - - M R N S P T F SEQ ID NO-1
128 - V K N S P T F SEQ ID NO-5
181 Q Q Y Q K T A E V A A A T S D T V K E K W N D M R N S S L F SEQ ID NO-6
148 K S F E E R V E T T V T S L K T K V G G T N P N G G S F E E SEQ ID NO-1
135 K S F E E K V E - - - - N L K S K V G G T K P A G G D F G E SEQ ID NO-5
211 K S F E S K L G S A L N N - - A K M A A S T S I - - - - - D SEQ ID NO-6
178 V L S S T A H A S A Q S L A G G S R R T K E E E L Q C SEQ ID NO-1
161 V L N S A A N A S A T T T E P L P E K T - Q E S L SEQ ID NO-5
234 H L A G A A R G P S Q T G T P V A E E A K P I S SEQ ID NO-6

FIGURE 5

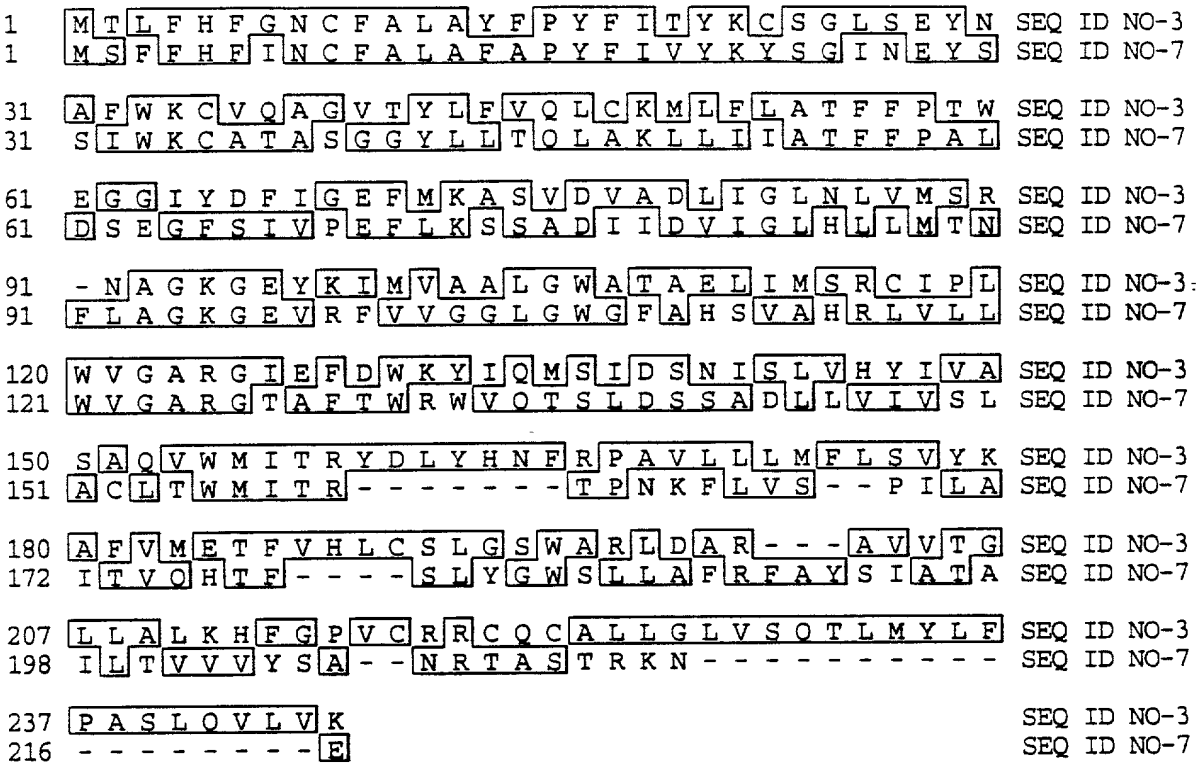
1 MTLFHFGNCFALAYFPYFITYKCSGLSEYN SEQ ID NO-3
1 MSFFHFINCFALAFAPYFIVYKYSGINEYS SEQ ID NO-7
31 AFWKCVQAGVTYLFVQLCKMLFLATFFPTW SEQ ID NO-3
31 SIWKCATASGGYLLTQLAKLLIIATFFPAL SEQ ID NO-7
61 EGGIYDFIGEFMKASVDVADLIGLNLVMSR SEQ ID NO-3
61 DSEGFSIVPEFLKSSADIIDVIGLHLLMTN SEQ ID NO-7
91 -NAGKGEYKIMVAALGWATAELIMSRCIPL SEQ ID NO-3
91 FLAGKGEVRFVVGGLGWGFAHSVAHRLVLL SEQ ID NO-7
120 WVGARGIEFDWKYIQMSIDSNISLVHYIVA SEQ ID NO-3
121 WVGARGTAFTWRWVQTSLDSSADLLVIVSL SEQ ID NO-7
150 SAQVWMITRYDLYHNFRPAVLLLMFLSVYK SEQ ID NO-3
151 ACLTWMITR-------TPNKFLVS--PILA SEQ ID NO-7
180 AFVMETFVHLCSLGSWARLDAR---AVVTG SEQ ID NO-3
172 ITVQHTF----SLYGWSLLAFRFAYSIATA SEQ ID NO-7
207 LLALKHFGPVCRRCQCALLGLVSQTLMYLF SEQ ID NO-3
198 ILTVVVYSA--NRTASTRKN---------- SEQ ID NO-7
237 PASLQVLVK SEQ ID NO-3
216 ---------E SEQ ID NO-7

FIGURE 6

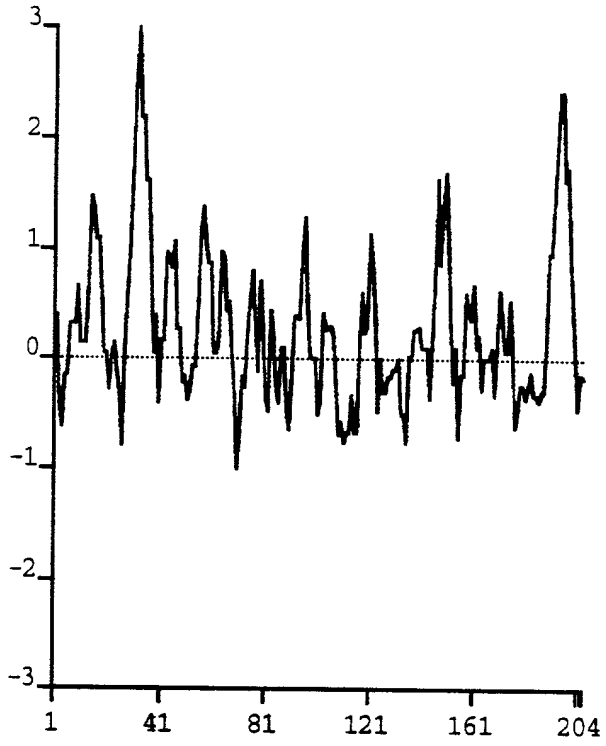

FIGURE 7

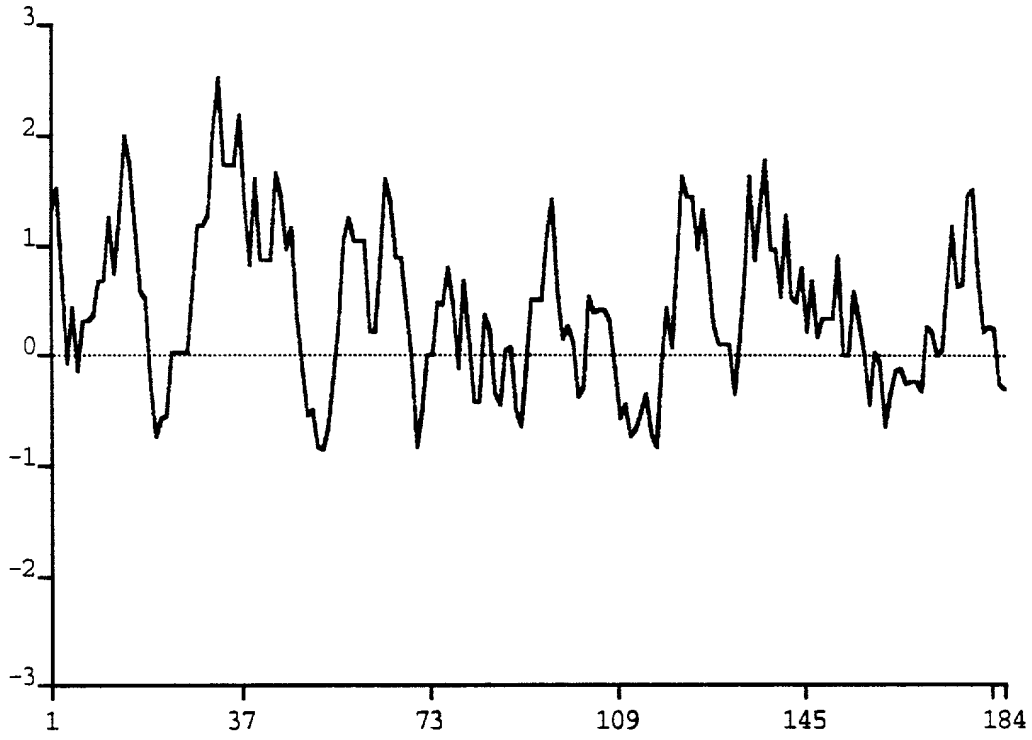

FIGURE 8

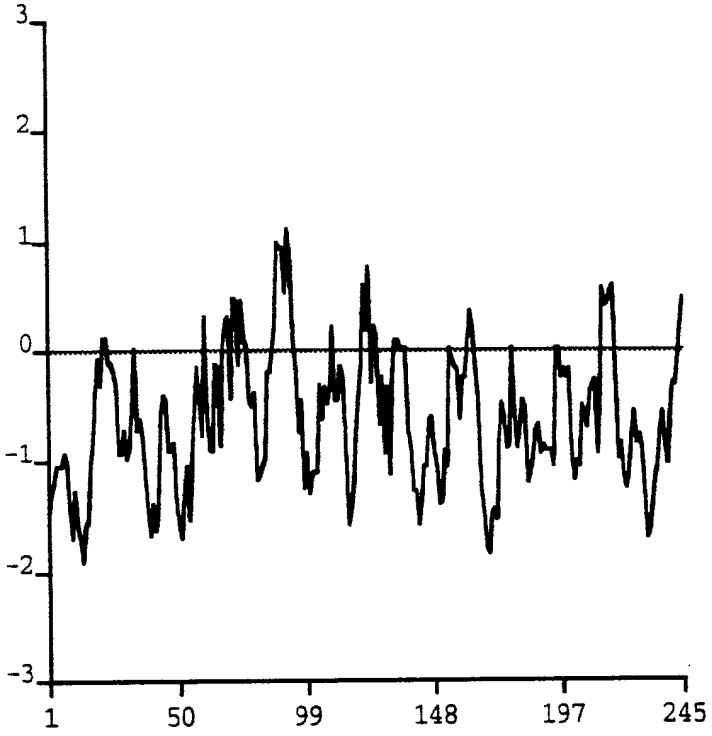

FIGURE 9

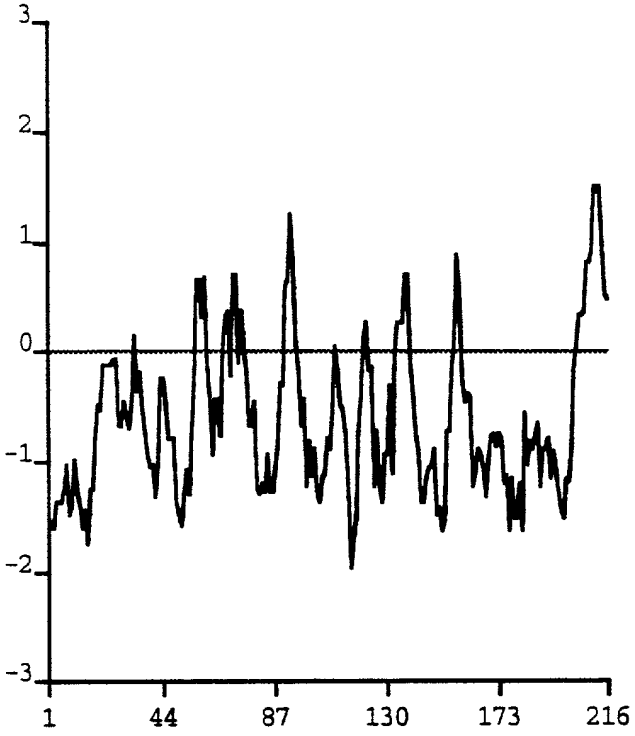

FIGURE 10

"Express Mail" mailing label number EM 062 546 944 US I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Assistant Commissioner for Patents, Box PATENT APPLICATION, Washington, D.C. 20231 on 29 Sep 98

By: _____

Printed: John J. Cherry

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Bandman et al.

Title: A NOVEL TUMOR PROTEIN

Serial No.:	To Be Assigned	Filing Date:	Herewith
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

Official Draftsman
Assistant Commissioner for Patents
Washington, D.C. 20231

SUBMISSION OF FORMAL DRAWINGS

Sir:

Transmitted herewith are Figures 1A, 1B, 1C, 2A, 2B, 2C, 3, 4A, 4B, 5, 6, 7, 8, 9, and 10, as fifteen (15) sheets of formal drawings for this application. Each sheet of drawing indicates the identifying indicia suggested in 37 C.F.R. 1.84(c) on the reverse side of the drawings.

Applicants believe that no fee is due with this paper. However, if the Commissioner determines that a fee is necessary, the Commissioner is hereby authorized to charge any additional fees associated with this communication or credit any overpayment to Deposit Account No. **09-0108. A duplicate copy of this communication is enclosed.**

If there are any questions regarding the above, the Examiner is invited to call the undersigned at 650-855-0555.

Respectfully submitted,

INCYTE PHARMACEUTICALS, INC.

Date: 29 September 98

Leanne C. Price
Reg. No. 42,090

3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 845-4166

"Express Mail" mailing label number EM 062 546 944 US I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Assistant Commissioner for Patents, Box PATENT APPLICATION, Washington, D.C. 20231 on 29 Sept 98

By: _____

Printed: John J. Cherry

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Bandman et al.

Title: A NOVEL TUMOR PROTEIN

Serial No.:	To Be Assigned	Filing Date:	Herewith
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

Official Draftsman
Assistant Commissioner for Patents
Washington, D.C. 20231

SUBMISSION OF FORMAL DRAWINGS

Sir:

Transmitted herewith are Figures 1A, 1B, 1C, 2A, 2B, 2C, 3, 4A, 4B, 5, 6, 7, 8, 9, and 10, as fifteen (15) sheets of formal drawings for this application. Each sheet of drawing indicates the identifying indicia suggested in 37 C.F.R. 1.84(c) on the reverse side of the drawings.

Applicants believe that no fee is due with this paper. However, if the Commissioner determines that a fee is necessary, the Commissioner is hereby authorized to charge any additional fees associated with this communication or credit any overpayment to Deposit Account No. **09-0108. A duplicate copy of this communication is enclosed.**

If there are any questions regarding the above, the Examiner is invited to call the undersigned at 650-855-0555.

Respectfully submitted,

INCYTE PHARMACEUTICALS, INC.

Date: 29 September 00

Leanne C. Price
Reg. No. 42,090

3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 845-4166

```
                9           18          27          36          45          54
5' GCC AGC TGC GTT CTG AGC CTG GGC GCA GCT ACC ATC TGC TCT GGG AAG CAC CAG

               63          72          81          90          99         108
   GGT GTC CCC GCC GCC CTC AGC TCG AAG TCA GCC ACC ATG GAG GCG CAG GCA CAA
                                                   M   E   A   Q   A   Q

              117         126         135         144         153         162
   GGT TTG TTG GAG ACT GAA CCG TTG CAA GGA ACA GAC GAA GAT GCA GTA GCC AGT
   G   L   L   E   T   E   P   L   Q   G   T   D   E   D   A   V   A   S

              171         180         189         198         207         216
   GCT GAC TTC TCT AGC ATG CTC TCT GAG GAG GAA AAG GAA GAG TTA AAA GCA GAG
   A   D   F   S   S   M   L   S   E   E   E   K   E   E   L   K   A   E

              225         234         243         252         261         270
   TTA GTT CAG CTA GAA GAC GAA ATT ACA ACA CTA CGA CAA GTT TTG TCA GCG AAA
   L   V   Q   L   E   D   E   I   T   T   L   R   Q   V   L   S   A   K

              279         288         297         306         315         324
   GAA AGG CAT CTA GTT GAG ATA AAA CAA AAA CTC GGC ATG AAC CTG ATG AAT GAA
   E   R   H   L   V   E   I   K   Q   K   L   G   M   N   L   M   N   E

              333         342         351         360         369         378
   TTA AAA CAG AAC TTC AGC AAA AGC TGG CAT GAC ATG CAG ACT ACC ACT GCC TAC
   L   K   Q   N   F   S   K   S   W   H   D   M   Q   T   T   T   A   Y
```

FIGURE 1A

```
        387         396         405         414         423         432
AAG AAA ACA CAT GAA ACC CTG AGT CAC GCA GGG CAA AAG GCA ACT GCA GCT TTC
K   K   T   H   E   T   L   S   H   A   G   Q   K   A   T   A   A   F

        441         450         459         468         477         486
AGC AAC GTT GGA ACG GCC ATC AGC AAG AAG TTC GGA GAC ATG AGT TAC TCC ATT
S   N   V   G   T   A   I   S   K   K   F   G   D   M   S   Y   S   I

        495         504         513         522         531         540
CGC CAT TCC ATA AGT ATG CCT GCT ATG AGG AAT TCT CCT ACT TTC AAA TCA TTT
R   H   S   I   S   M   P   A   M   R   N   S   P   T   F   K   S   F

        549         558         567         576         585         594
GAG GAG AGG GTT GAG ACA ACT GTC ACA AGC CTC AAG ACG AAA GTA GGC GGT ACG
E   E   R   V   E   T   T   V   T   S   L   K   T   K   V   G   G   T

        603         612         621         630         639         648
AAC CCT AAT GGA GGC AGT TTT GAG GAG GTC CTC AGC TCC ACG GCC CAT GCC AGT
N   P   N   G   G   S   F   E   E   V   L   S   S   T   A   H   A   S

        657         666         675         684         693         702
GCC CAG AGC TTG GCA GGA GGC TCC CGG CGG ACC AAG GAG GAG GAG CTG CAG TGC
A   Q   S   L   A   G   G   S   R   R   T   K   E   E   E   L   Q   C

        711         720         729         738         747         756
TAA GTC CAG CCA GCG TGC AGT GCA TCC AGA AAC CGG CCA CTA CCC AGC CCA TCT
```

FIGURE 1B

```
          765          774          783
NTG CCT GTG CTT ATC CAG ATA AGA AGA CCA AA 3'
```

FIGURE 1C

```
                9          18          27          36          45          54
5' TMG MKC GCG GGC CCC CGC CAG TCA GGT GGG TGC CAG GCC CTG GCC GTG GCG AAA

               63          72          81          90          99         108
GAG CCG GCG GAG GGA GGA CCC GCT CCC GGA GAC GCC GCC TCG CGA TCC CCG CGC

              117         126         135         144         153         162
GGG CGG GAC CGG GCG GCC GGC ATC ATG ACC CTG TTT CAC TTC GGG AAC TGC TTC
                                    M   T   L   F   H   F   G   N   C   F

              171         180         189         198         207         216
GCT CTT GCC TAC TTC CCC TAC TTC ATC ACC TAC AAG TGC AGC GGC CTG TCC GAG
A   L   A   Y   F   P   Y   F   I   T   Y   K   C   S   G   L   S   E

              225         234         243         252         261         270
TAC AAC GCC TTC TGG AAA TGC GTC CAG GCT GGA GTC ACC TAC CTC TTT GTC CAA
Y   N   A   F   W   K   C   V   Q   A   G   V   T   Y   L   F   V   Q

              279         288         297         306         315         324
CTC TGC AAG ATG CTG TTC TTG GCC ACT TTC TTT CCC ACC TGG GAA GGC GGC ATC
L   C   K   M   L   F   L   A   T   F   F   P   T   W   E   G   G   I

              333         342         351         360         369         378
TAT GAC TTC ATT GGG GAG TTC ATG AAG GCC AGC GTG GAT GTG GCA GAC CTG ATA
Y   D   F   I   G   E   F   M   K   A   S   V   D   V   A   D   L   I
```

FIGURE 2A

```
          387         396         405         414         423         432
GGT CTA AAC CTT GTC ATG TCC CGG AAT GCC GGC AAG GGA GAG TAC AAG ATC ATG
G   L   N   L   V   M   S   R   N   A   G   K   G   E   Y   K   I   M

          441         450         459         468         477         486
GTT GCT GCC CTG GGC TGG GCC ACT GCT GAG CTT ATT ATG TCC CGC TGC ATT CCC
V   A   A   L   G   W   A   T   A   E   L   I   M   S   R   C   I   P

          495         504         513         522         531         540
CTA TGG GTC GGA GCC CGG GGC ATT GAG TTT GAC TGG AAG TAC ATC CAG ATG AGC
L   W   V   G   A   R   G   I   E   F   D   W   K   Y   I   Q   M   S

          549         558         567         576         585         594
ATA GAC TCC AAC ATC AGT CTG GTC CAT TAC ATC GTC GCG TCT GCT CAG GTC TGG
I   D   S   N   I   S   L   V   H   Y   I   V   A   S   A   Q   V   W

          603         612         621         630         639         648
ATG ATA ACA CGC TAT GAT CTG TAC CAC AAC TTC CGG CCA GCT GTC CTT CTG CTG
M   I   T   R   Y   D   L   Y   H   N   F   R   P   A   V   L   L   L

          657         666         675         684         693         702
ATG TTC CTC AGT GTC TAC AAG GCC TTT GTT ATG GAG ACC TTC GTC CAC CTC TGC
M   F   L   S   V   Y   K   A   F   V   M   E   T   F   V   H   L   C

          711         720         729         738         747         756
TCG CTG GGC AGT TGG GCA RCT CTA MTG GCC CGA GCA GTG GTA ACG GGG CTG CTG
S   L   G   S   W   A   X   L   X   A   R   A   V   V   T   G   L   L
```

FIGURE 2B

```
          765          774          783          792          801          810
GCC CTC AAG CAC TTT GGS CCT GTA TGT CGS CGT TGT CAA TGT GCA CTY CTA GGC
A   L   K   H   F   G   P   V   C   R   R   C   Q   C   A   L   L   G

          819          828          837          846          855          864
TTG GTG TCT CAG ACA TTG ATG TAC CTT TTC CCT GCC TCA CTC CAG GTT TTA GTG
L   V   S   Q   T   L   M   Y   L   F   P   A   S   L   Q   V   L   V

          873          882
AAG TAA ACA GTA TTT GGA AAG TT 3'
K
```

FIGURE 2C

Library	Lib Description	Abun	Pct Abun
BRAINOM02	brain, 55 M, NORM, WM	1	0.0454
UTRSNOT01	uterus, 59 F	1	0.0393
TLYMNOR01	lymphocytes (non-adher PBMNC), 24 M, RP	1	0.0372
BRSTNOT02	breast, 55 F, match to BRSTTUT01	2	0.0317
PROSTUT08	prostate tumor, 60 M, match to PROSNOT14	1	0.0266
PROSNOT14	prostate, 60 M, match to PROSTUT08	1	0.0256
LIVRNOM01	liver, 49 M, WM	1	0.0254
PROSNOT15	prostate, 66 M, match to PROSTUT10	1	0.0241
NERVMSM01	multiple sclerosis, 46 M, NORM, WM	1	0.0228
HNT2AGT01	hNT2 cell line, post-mitotic neurons	1	0.0190
BRAITUT02	brain tumor, metastasis, 58 M	1	0.0169
LIVSFEM02	liver/spleen, fetal M, NORM, WM	2	0.0053

Electronic Northern Results returned a total of 12 row(s).

FIGURE 3

Library	Lib Description	Abun	Pct Abun
COLNNOT22	colon, 56 F	2	0.0554
COLNPOT01	colon polyp, 40 F	2	0.0513
PROSNOT18	prostate, hyperplasia, 58 M	2	0.0513
MUSCNOT02	muscle, psoas, 12 M	1	0.0382
STOMTUT01	stomach tumor, 52 M, match to STOMNOT02	1	0.0368
SINTNOT02	small intestine, 55 F	1	0.0337
LVENNOT03	heart, left ventricle, 31 M	1	0.0337
MMLR3DT01	macrophages (adher PBMNC), M/F, 72-hr MLR	1	0.0331
PROSTUT01	prostate tumor, 50 M, match to PROSNOT02	1	0.0310
LUNGTUT03	lung tumor, 69 M, match to LUNGNOT15	1	0.0308
BLADTUT02	bladder tumor, 80 F, match to BLADNOT03	1	0.0305
BRAITUT08	brain tumor, astrocytoma, 47 M	2	0.0293
PROSTUT12	prostate tumor, 65 M, match to PROSNOT20	1	0.0279
BLADNOT04	bladder, 28 M	1	0.0278
TESTTUT02	testicular tumor, 31 M	1	0.0278
THYRNOT03	thyroid tumor, adenoma, 28 F	2	0.0277
SINTNOT13	small intestine, ileum, ulcerative cholitis, 25 F	1	0.0275
COLNTUT03	colon tumor, 62 M, match to COLNNOT16	1	0.0272
BLADTUT05	bladder tumor, 66 M, match to BLADNOT06	1	0.0268
KIDNTUT01	kidney tumor, Wilms, 8m F	1	0.0267
PENITUT01	penis tumor, carcinoma, 64 M	1	0.0267
COLNNOT23	colon, 16 M	1	0.0264
BRAITUT13	brain tumor, meningioma, 68 M	1	0.0262

FIGURE 4A

LIVRTUT01	liver tumor, metastasis, 51 F	1	0.0259
PROSNOT14	prostate, 60 M, match to PROSTUT08	1	0.0256
BRSTTUT08	breast tumor, 45 F, match to BRSTNOT09	1	0.0254
BMARNOT03	bone marrow, 16 M	1	0.0242
RATRNOT02	heart, right atrium, 39 M	1	0.0236
PANCNOT01	pancreas, 29 M	1	0.0214
LUNGNOT04	lung, 2 M	1	0.0183
SYNORAT04	synovium, wrist, rheumatoid, 62 F	1	0.0174
PLACNOT02	placenta, fetal F	1	0.0168
BRSTNOT03	breast, 54 F, match to BRSTTUT02	1	0.0147
SPLNNOT04	spleen, 2 M	1	0.0128
PROSNOT06	prostate, 57 M, match to PROSTUT04	1	0.0114
LUNGFET03	lung, fetal F	1	0.0091

Electronic Northern Results returned a total of 36 row(s).

FIGURE 4B

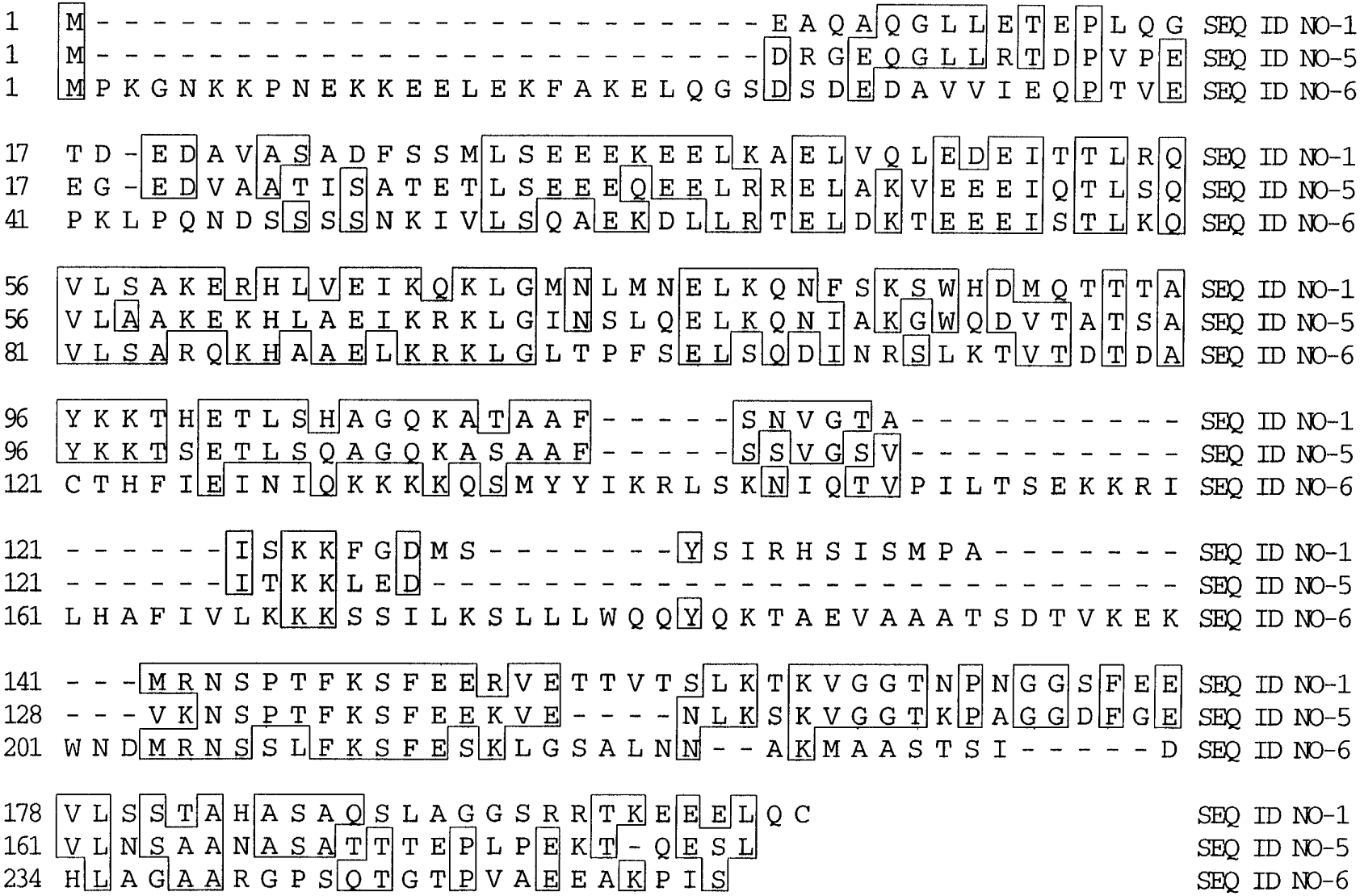
1 M - E A Q A Q G L L E T E P L Q G SEQ ID NO-1
1 M - D R G E Q G L L R T D P V P E SEQ ID NO-5
1 M P K G N K K P N E K K E E L E K F A K E L Q G S D S D E D A V V I E Q P T V E SEQ ID NO-6
17 T D - E D A V A S A D F S S M L S E E E K E E L K A E L V Q L E D E I T T L R Q SEQ ID NO-1
17 E G - E D V A A T I S A T E T L S E E E Q E E L R R E L A K V E E E I Q T L S Q SEQ ID NO-5
41 P K L P Q N D S S S S N K I V L S Q A E K D L L R T E L D K T E E E I S T L K Q SEQ ID NO-6
56 V L S A K E R H L V E I K Q K L G M N L M N E L K Q N F S K S W H D M Q T T T A SEQ ID NO-1
56 V L A A K E K H L A E I K R K L G I N S L Q E L K Q N I A K G W Q D V T A T S A SEQ ID NO-5
81 V L S A R Q K H A A E L K R K L G L T P F S E L S Q D I N R S L K T V T D T D A SEQ ID NO-6
96 Y K K T H E T L S H A G Q K A T A A F - - - - - S N V G T A - - - - - - - - - - - SEQ ID NO-1
96 Y K K T S E T L S Q A G Q K A S A A F - - - - - S S V G S V - - - - - - - - - - - SEQ ID NO-5
121 C T H F I E I N I Q K K K K Q S M Y Y I K R L S K N I Q T V P I L T S E K K R I SEQ ID NO-6
121 - - - - - - - I S K K F G D M S - - - - - - - Y S I R H S I S M P A - - - - - - - SEQ ID NO-1
121 - - - - - - - I T K K L E D - SEQ ID NO-5
161 L H A F I V L K K K S S I L K S L L L W Q Q Y Q K T A E V A A A T S D T V K E K SEQ ID NO-6
141 - - - M R N S P T F K S F E E R V E T T V T S L K T K V G G T N P N G G S F E E SEQ ID NO-1
128 - - - V K N S P T F K S F E E K V E - - - - N L K S K V G G T K P A G G D F G E SEQ ID NO-5
201 W N D M R N S S L F K S F E S K L G S A L N N - - A K M A A S T S I - - - - - D SEQ ID NO-6
178 V L S S T A H A S A Q S L A G G S R R T K E E E L Q C SEQ ID NO-1
161 V L N S A A N A S A T T T E P L P E K T - Q E S L SEQ ID NO-5
234 H L A G A A R G P S Q T G T P V A E E A K P I S SEQ ID NO-6

FIGURE 5

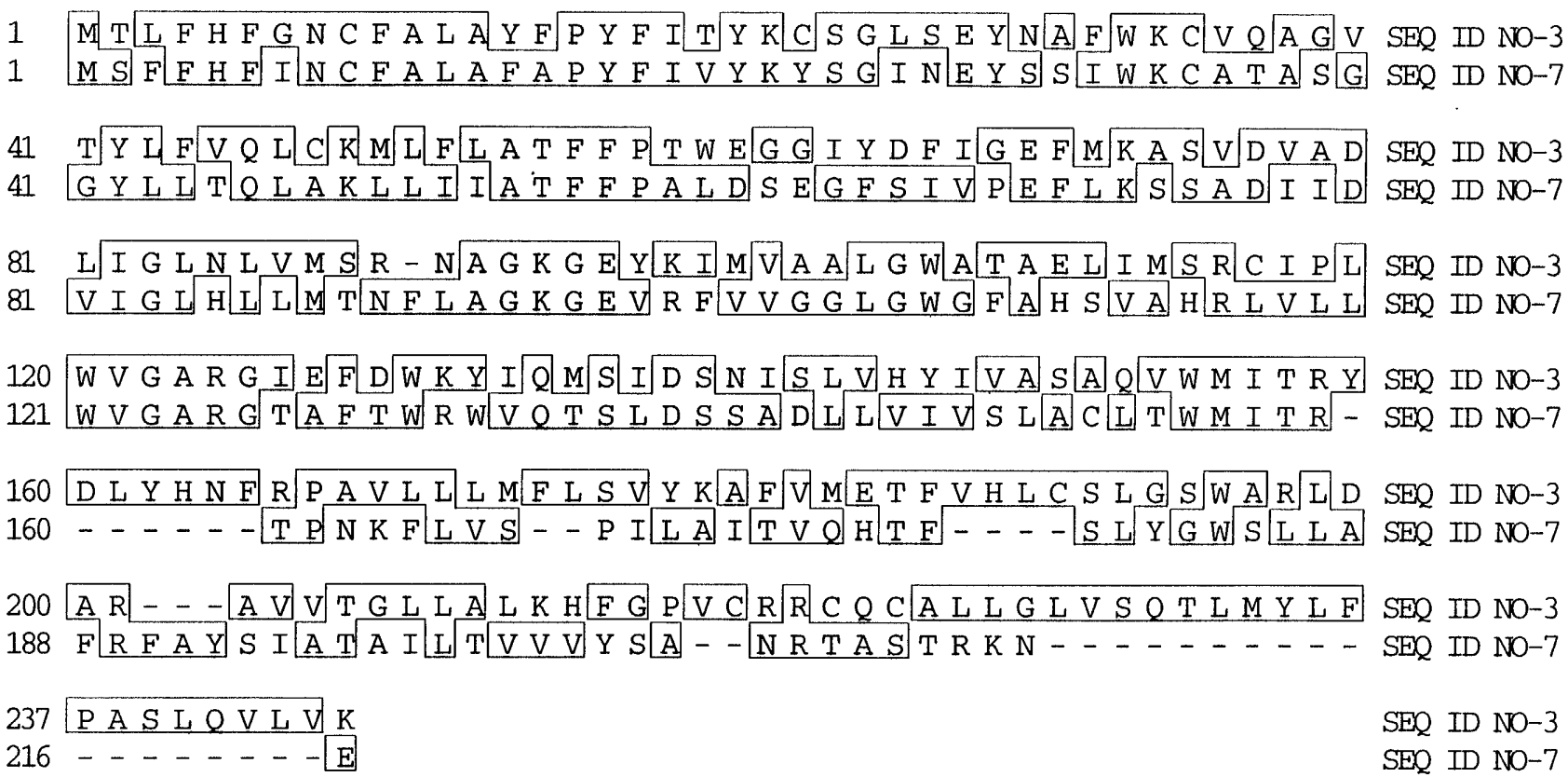
1 MTLFHFGNCFALAYFPYFITYKCSGLSEYNAFWKCVQAGV SEQ ID NO-3
1 MSFFHFINCFALAFAPYFIVYKYSGINEYSSIWKCATASG SEQ ID NO-7
41 TYLFVQLCKMLFLATFFPTWEGGIYDFIGEFMKASVDVAD SEQ ID NO-3
41 GYLLTQLAKLLIIATFFPALDSEGFSIVPEFLKSSADIID SEQ ID NO-7
81 LIGLNLVMSR-NAGKGEYKIMVAALGWATAELIMSRCIPL SEQ ID NO-3
81 VIGLHLLMTNFLAGKGEVRFVVGGLGWGFAHSVAHRLVLL SEQ ID NO-7
120 WVGARGIEFDWKYIQMSIDSNISLVHYIVASAQVWMITRY SEQ ID NO-3
121 WVGARGTAFTWRWVQTSLDSSADLLVIVSLACLTWMITR- SEQ ID NO-7
160 DLYHNFRPAVLLLMFLSVYKAFVMETFVHLCSLGSWARLD SEQ ID NO-3
160 -------TPNKFLVS--PILAITVQHTF----SLYGWSLLA SEQ ID NO-7
200 AR---AVVTGLLALKHFGPVCRRCQCALLGLVSQTLMYLF SEQ ID NO-3
188 FRFAYSIATAILTVVVYSA--NRTASTRKN---------- SEQ ID NO-7
237 PASLQVLVK SEQ ID NO-3
216 ---------E SEQ ID NO-7

FIGURE 6

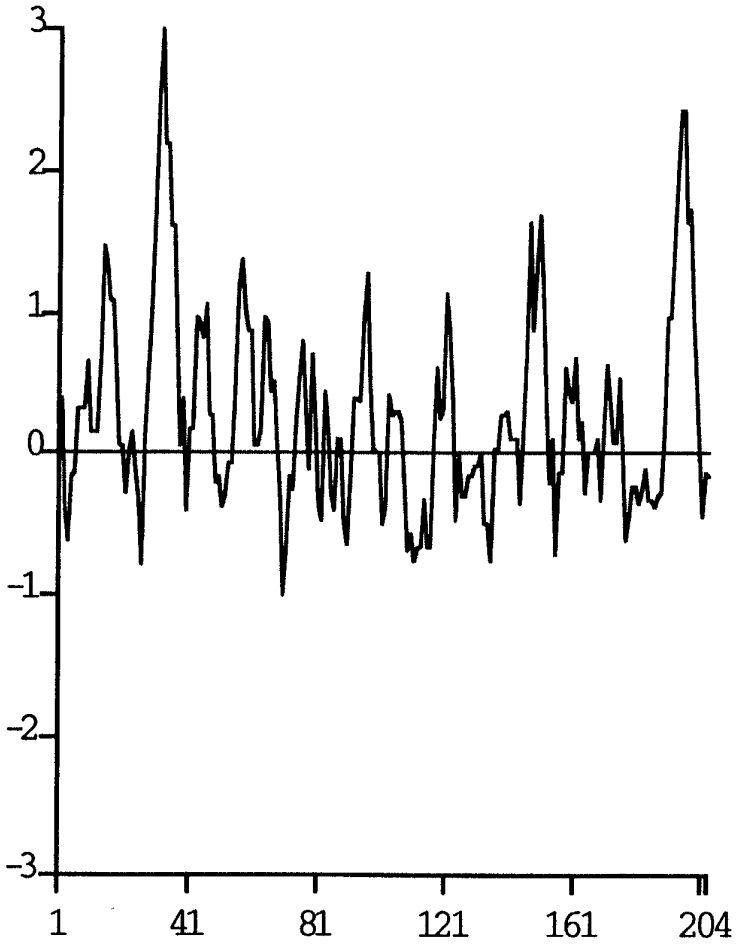

FIGURE 7

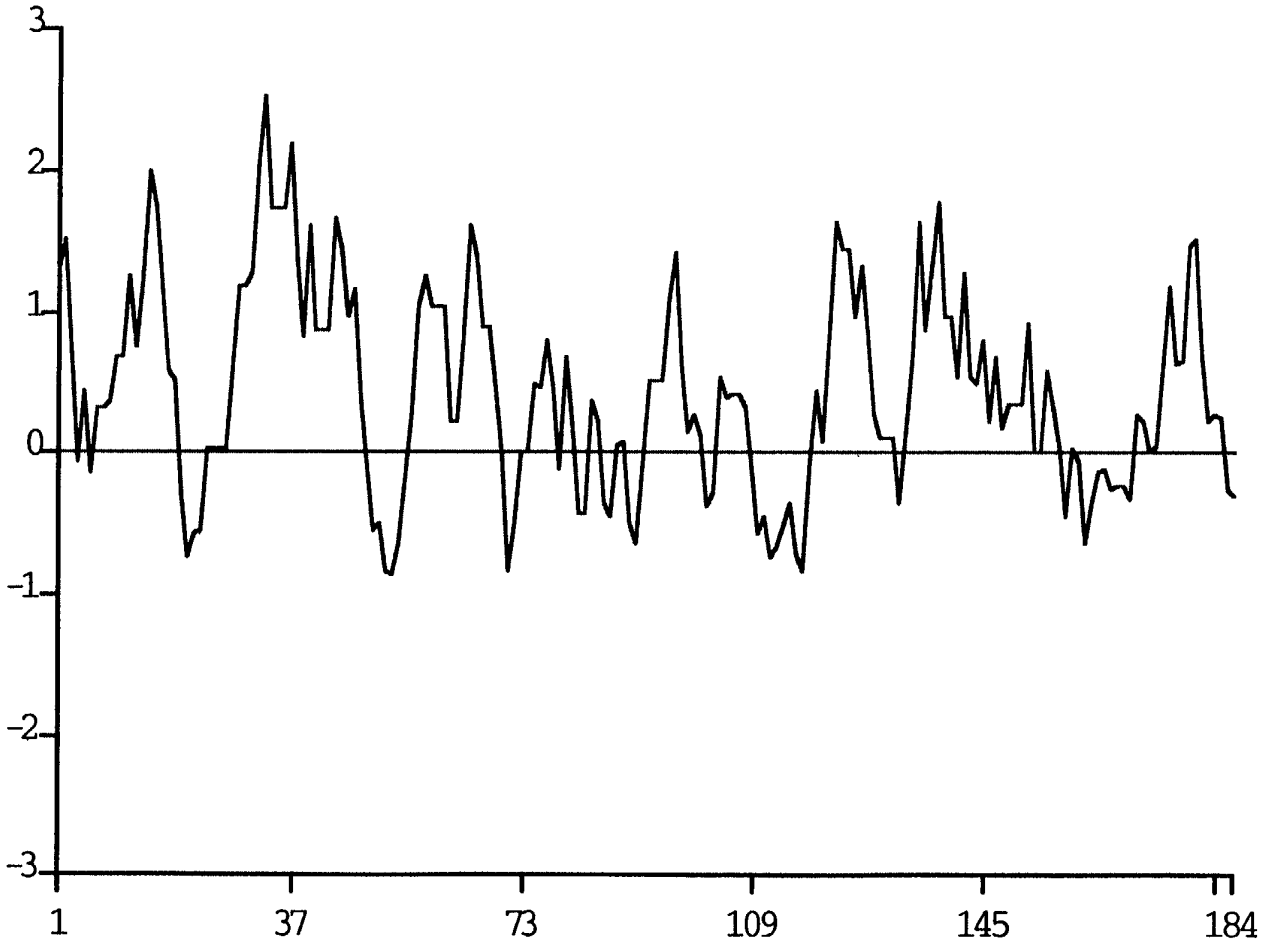

FIGURE 8

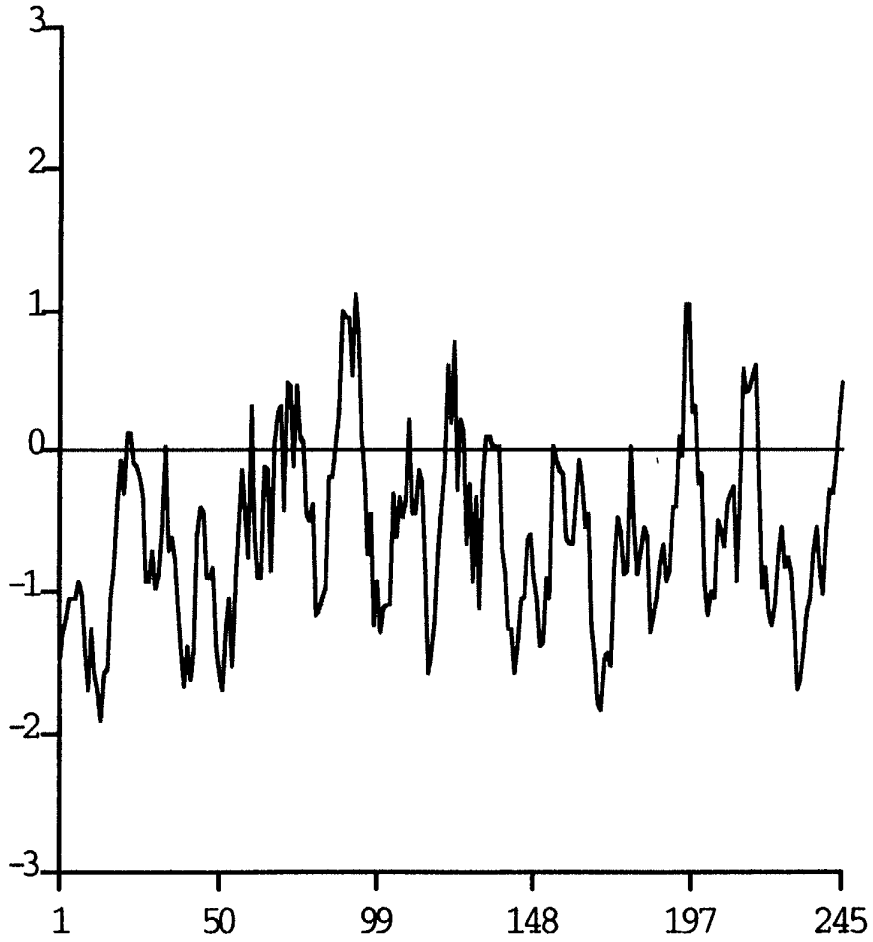

FIGURE 9

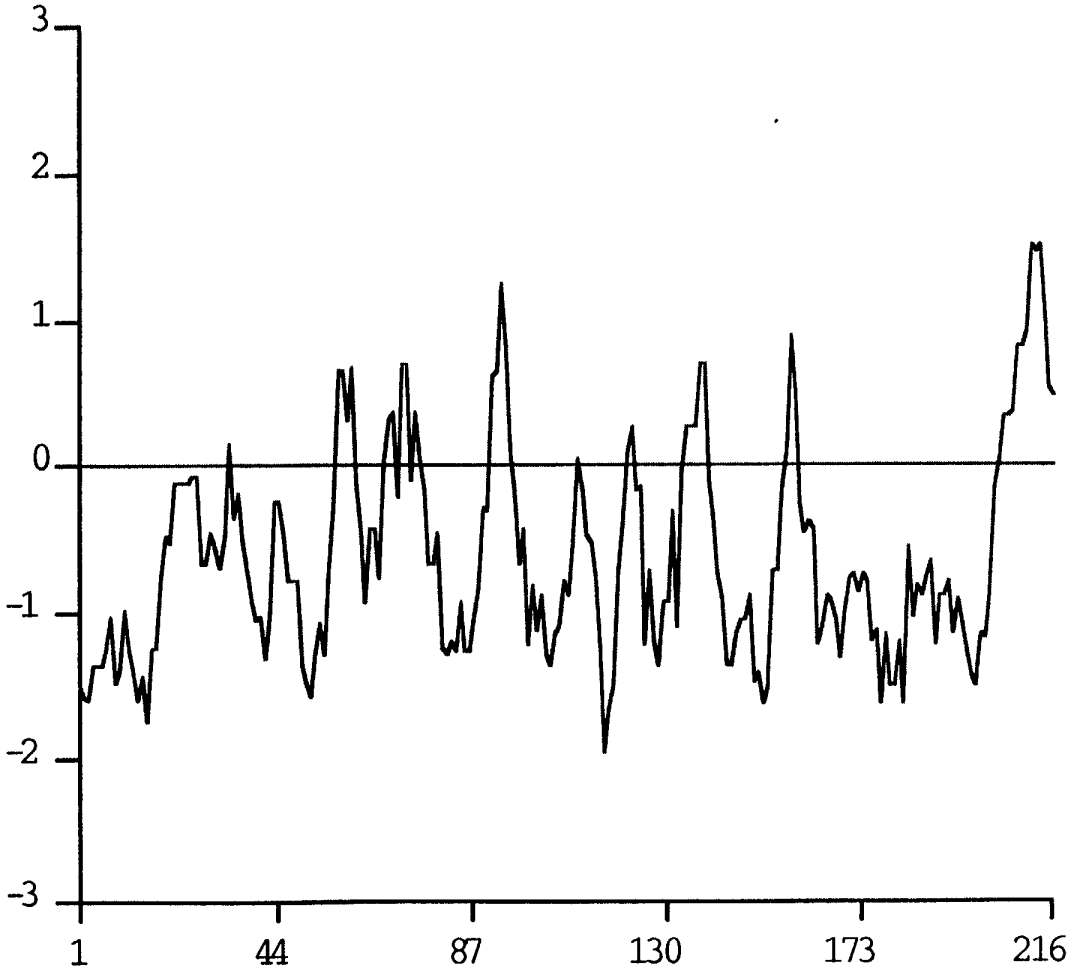

FIGURE 10

Docket # PF-0126 US

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

A NOVEL TUMOR PROTEIN

the specification of which:

/__/ is attached hereto.

/ X / was filed on September 18, 1996, as application Serial No. 08/715,204 and if this box contains an X /___/, was amended on _____.

/_/ was filed as Patent Cooperation Treaty international application No. _____ on _____, 19__, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____19_, and if this box contains an X /_/, was amended on _____ _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

LUCY J. BILLINGS **Registration No. 36,749**
MICHAEL C. CERRONE **Registration No. 39,132**

respectively and individually, as my attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

LUCY J. BILLINGS, ESQ.
INCYTE PHARMACEUTICALS, INC.
3174 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 415-855-0555 FAX: 415-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United

States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

**IMPORTANT: Before this declaration is signed, the patent application (the specification, the claims and this declaration) must be read and understood by each person signing it, and no changes may be made in the application after this declaration has been signed.*

Sole Inventor or First Joint Inventor:

Full name: **OLGA BANDMAN**

Signature: Olga Bandman

Date: Jan 6 , 1997 013

Citizenship: United States of America

Residence: Mountain View, California

P.O. Address: 366 Anna Avenue
Mountain View, CA 94043

Second Joint Inventor:

Full name: **JANICE AU-YOUNG**

Signature: Janice Au-Young

Date: January 6 , 1997

Citizenship: United States of America

Residence: Berkeley, California

P.O. Address: 1419 Kains Avenue
Berkeley, CA 94702

Third Joint Inventor:

Full name: **SURYA K. GOLI**

Signature: Surya Kumari Goli

Date: 1/6/1997 GSK

Citizenship: India

Residence: Sunnyvale, California

P.O. Address: 620 Iris Avenue #338
Sunnyvale, CA 94086

Fourth Joint Inventor:

Full name: **JENNIFER L. HILLMAN**

Signature: Jennifer L. Hillman

Date: Jan 2. 1997 JLH

Citizenship: United States of America

Residence: San Jose, California

P.O. Address: 467 North Third Street #3
San Jose, CA 95050

"Express Mail" mailing label number EM 062 546 944 US I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Assistant Commissioner for Patents, Box PATENT APPLICATION, Washington, D.C. 20231 on 29 Sept. 98

By: _____ Printed: John J. Cherry

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Bandman et al.

Title:	A NOVEL TUMOR PROTEIN		
Serial No.:	To Be Assigned	Filing Date:	Herewith
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

Assistant Commissioner for Patents
Washington, D.C. 20231

ASSOCIATE POWER OF ATTORNEY

Sir:

I hereby appoint the following attorneys, whose post office address is 3174 Porter Drive, Palo Alto, California 94304, as associate attorneys in the above-entitled application, to prosecute this application, to make alterations and amendments therein, and to transact all business in the Patent and Trademark Office connected therewith:

Leanne C. Price	**Registration No. 42,090**
Lynn E. Murry	**Registration No. 42,918**
David G. Streeter	**Registration No. 43,168**

Please continue to address all future communications to:

Legal Department
Incyte Pharmaceuticals, Inc.
3174 Porter Drive
Palo Alto, California 94304

Respectfully submitted,
INCYTE PHARMACEUTICALS, INC.

Date: Sept 28 1998

Michael C. Cerrone, Ph.D.
Reg. No. 39,132

3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 845-4166

COPY

Docket No.: PF-0126 US

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Assistant Commissioner for Patents, Box Issue Fee, Washington, D.C. 20231 on 5/15/98.

By: Peggy Fazzino

Printed: Peggy Fazzino

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Bandman et al.

Title:	A NOVEL TUMOR PROTEIN		
Serial No.:	08/715,204	Filing Date:	September 17, 1996
Examiner:	Johnson, N.	Group Art Unit:	1642
		Batch No.:	A61

Assistant Commissioner for Patents
Box Issue Fee
Washington, D.C. 20231

ASSOCIATE POWER OF ATTORNEY

Sir:

I hereby appoint the following attorneys, whose post office address is 3174 Porter Drive, Palo Alto, California 94304, as associate attorneys in the above-entitled application, to prosecute this application, to make alterations and amendments therein, and to transact all business in the Patent and Trademark Office connected therewith:

Sheela Mohan-Peterson	**Registration No. 41,201**
Colette C. Muenzen	**Registration No. 39,784**
Karen J. Zeller	**Registration No. 37,071**

Please continue to address all future communications to:

Legal Department
Incyte Pharmaceuticals, Inc.
3174 Porter Drive
Palo Alto, California 94304

Respectfully submitted,
INCYTE PHARMACEUTICALS, INC.

Date: March 27, 1998

Lucy J. Billings
Lucy J. Billings
Reg. No. 36,749

3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 845-4166